

МИНИСТЕРСТВО НАУКИ И ВЫСШЕГО ОБРАЗОВАНИЯ РОССИЙСКОЙ ФЕДЕРАЦИИ

РОССИЙСКОЕ НАУЧНО-ТЕХНИЧЕСКОЕ ОБЩЕСТВО РАДИОТЕХНИКИ,
ЭЛЕКТРОНИКИ И СВЯЗИ им. А.С.ПОПОВА

ФЕДЕРАЛЬНОЕ ГОСУДАРСТВЕННОЕ БЮДЖЕТНОЕ ОБРАЗОВАТЕЛЬНОЕ
УЧРЕЖДЕНИЕ ВЫСШЕГО ОБРАЗОВАНИЯ
«НИЖЕГОРОДСКИЙ ГОСУДАРСТВЕННЫЙ ТЕХНИЧЕСКИЙ УНИВЕРСИТЕТ
им. Р.Е. АЛЕКСЕЕВА»

ИНСТИТУТ РАДИОЭЛЕКТРОНИКИ И ИНФОРМАЦИОННЫХ ТЕХНОЛОГИЙ



«ИНФОРМАЦИОННЫЕ СИСТЕМЫ И ТЕХНОЛОГИИ» ИСТ–2021

***Сборник материалов XXVII Международной
научно-технической конференции***

© Нижегородский государственный
технический университет
им. Р.Е. Алексева, 2021

Нижний Новгород, 2021

УДК 621:681
ББК 32.97
И650

Информационные системы и технологии - 2021: [Электронный ресурс]: сборник материалов XXVII Международной научно-технической конференции – Электрон. дан. – Н. Новгород: Нижегород. гос. техн. ун-т им. Р.Е. Алексеева, 2021. – 1 электрон. диск (CD-ROM): зв., цв., 12 см. – Систем. требования: ПК с процессором 486; ОЗУ 8 Мб; операц/система Windows 95; CD-ROM дисковод; мышь. – Загл. с экрана. – 200 экз.

В сборнике представлены материалы докладов XXVII Международной научно-технической конференции ИСТ-2021, проведенной 23, 24 апреля 2021 г. дирекцией Института радиоэлектроники и информационных технологий при поддержке, оказанной ректоратом НГТУ им. Р.Е. Алексеева и Нижегородским областным правлением РНТОРЭС им. А.С. Попова.

Конференция проведена в смешанном режиме: очно и с дистанционным доступом в формате вебинара на специализированной электронной платформе в сети Интернет.

Публикуемые материалы представляют тематику, круг научных интересов и состояние исследований представителей научных и высших учебных заведений Белоруссии, Украины и 13 городов РФ - преподавателей, научных сотрудников, докторантов, аспирантов, магистрантов и студентов - участников НИРС, а также сотрудников МГУ им. М.В. Ломоносова, МГТУ им. Н.Э. Баумана, МЭИ, МГТУ «СТАНКИН», Высшей школы экономики, Владимирского государственного университета, Нижегородского государственного технического университета им. Р.Е. Алексеева, Нижегородского государственного университета им. Н.И. Лобачевского, Волжского государственного университета водного транспорта, Волго-Вятского филиала МТУСИ; представителей более 25 предприятий и НИИ: ИПУ РАН, РФЯЦ-ВНИИЭФ, ИПФ РАН, ИФМ РАН, НИИИС им. Ю.Е. Седакова, НИФТИ, НПП «Полет», АПЗ им. П.И. Пландина, АПКБ и других организаций.

Организационный комитет:

А.В. Мякинков (председатель), В.П. Хранилов (зам. председателя),
А.А. Куркин, Н.Ю. Бабанов, В.Г. Баранов, О.Г. Берестнева, Д.В. Жевнерчук,
В.Ю. Карпычев, В.В. Кондратьев, В.Р. Милов, Д.А. Ляхманов, С.Л. Моругин,
А.С. Раевский, А.Г. Рындык, Э.С. Соколова, Ю.М. Туляков, М.В. Ульянов,
А.Д. Филинских, В.Л. Ягодкин

ISBN 978-5-502-01481-6

© Нижегородский государственный
технический университет им. Р.Е. Алексеева, 2021

СОДЕРЖАНИЕ

<i>Пленарное заседание</i>	4
<i>Секция 1</i> Телекоммуникационные системы	29
<i>Секция 2</i> Информационные технологии (Информационное обеспечение систем автоматизации промышленных объектов).....	239
<i>Секция 3</i> Информационные технологии (Автоматизация проектирования).....	340
<i>Секция 4</i> Информационная безопасность систем и сетей.....	443
<i>Секция 5.1</i> Техническая кибернетика (Системы обработки информации)	546
<i>Секция 5.2</i> Техническая кибернетика (Интеллектуальные системы управления).....	619
<i>Секция 5.3</i> Техническая кибернетика (Информационное моделирование когнитивных процессов).....	730
<i>Секция 6</i> Математическое моделирование геофизических процессов	848
<i>Секция 7</i> Радиотехнические системы и устройства.....	968
<i>Секция 8</i> Проектирование радиоэлектронных устройств ВЧ- и СВЧ- диапазонов	1082

В.Г. Баранов, В.П. Хранилов

ГЛЯДЯ В ПРОШЛОЕ, НАДО СТРОИТЬ БУДУЩЕЕ

Нижегородский государственный технический университет им. Р.Е. Алексеева

Представлена история создания и развития радиотехнического факультета в составе Нижегородского политехнического института (НГТУ им. Р.Е. Алексеева). Описаны структура и состав института радиоэлектроники и информационных технологий. Представлена информация о ведущих научно-педагогических сотрудниках. Даны сведения о наградах выдающихся сотрудников и выпускников.

Ключевые слова: Нижегородский политехнический институт, радиотехнический факультет, Горьковский индустриальный институт, факультет радиоэлектроники и технической кибернетики, факультет информационных систем и технологий, институт радиоэлектроники и информационных технологий.

История становления

В январе 2021 года исполнилось 85 лет со дня создания радиотехнического факультета Горьковского индустриального института – ныне института радиоэлектроники и информационных технологий НГТУ им. Р.Е. Алексеева.

Организация факультета – веление времени, предъявившего высокие и жесткие требования к инженерным кадрам в канун Великой отечественной войны, была логичным шагом создания специализированной структуры по подготовке радиоинженеров.

Основания для такого решения достаточно очевидны: во-первых, в Н. Новгороде на базе радиотелефонного завода («Siemens & Halske»), открытого в 1917 г. и созданной в 1918 г. Нижегородской радиолaborатории началось развитие радиопромышленности, требовавшей инженеров и техников; во-вторых, сначала в Нижегородском государственном университете на механическом факультете, а в 1930 г., после расформирования НГУ, в Нижегородском механико-машиностроительном институте была организована “штучная” подготовка инженеров по радиотехнике и был накоплен опыт преподавания ряда радиотехнических дисциплин; в-третьих, на базе центральной военно-индустриальной радиолaborатории (ЦВИРЛ), созданной в 1929 г. на основе переведенного из Ленинграда военного отдела центральной радиолaborатории (ЦРЛ), были открыты высшие вечерние двухгодичные курсы для инженерно-технических работников (в основном ЦВИРЛ и радиотелефонного завода) с началом занятий с 1.10.1930 г.; в-четвертых, в Н. Новгороде в течение продолжительного времени успешно работали выдающиеся ученые-радиотехники А.С. Попов (на Нижегородской ярмарке), М.А. Бонч-Бруевич, В.П. Вологдин, А.Ф. Шорин, А.Н. Кугушев, В.М. Лещинский, В.К. Лебединский, В.В. Татарин, П.А. Остряков, О.В. Лосев.

Первым деканом факультета и заведующим кафедрой радиотехники был назначен профессор Р.В. Львович.

Рудольф Владимирович Львович (1878-1937 г.г.) закончил в 1903 г. Рижский политехнический институт по специальности “Инженер-технолог” и электротехнический институт при университете в Льеже (Бельгия). Работал в Одесском политехническом институте и в ЦВИРЛ научным руководителем по передатчикам.

Заведующим кафедрой электротехники был назначен доцент В.Г. Власов.

Первый выпуск состоялся уже в 1939 году. Столь “скорый” выпуск инженеров обусловлен переводом в 1936г. группы студентов ЛЭИС на 3-й курс радиофака.

Председателем ГЭК был В.А. Авдентов – известный радиоинженер и педагог, позже лауреат Государственной премии, один из руководителей КБ ГТЗ им. В.И. Ленина (ныне ФНПЦ “ОАО “НИИРТ”).

Среди первых выпускников, а их было шестнадцать, Курячев В.П. – лауреат Государственной премии, один из организаторов НИИИС им. Ю.Е. Седакова, Бочкарев Б.А. – разработчик базовой технологии резисторов МЛТ, лауреат Государственной премии, Увяткин Б.Д. – лауреат Государственной премии в области радиосвязи, крупный организатор радиопромышленности, директор радиозаводов в ряде городов России, Мальцев Б.А. – лауреат Государственной премии за разработку самолетного радиокompасного приемника.

В 1938г. факультет был преобразован в “Спецфак” с правилами приема, установленными для втузов народного комиссариата оборонной промышленности и сроком обучения 5,5 лет.

К сожалению, в 1941 году не все студенты защитили дипломные проекты - грянула война и многие студенты и преподаватели добровольно ушли на фронт. Студент Ю.И. Онусайтис стал Героем Советского Союза. Впоследствии Юрий Иосифович – генерал-майор, заведовал кафедрой Военной академии и долгие годы поддерживал связь с политехом.

В 1947г. по инициативе декана И.В. Типашова “Спецфак” преобразовали в электротехнический факультет, где наряду со специальностью “Радиотехника” (прием – 3 группы) была открыта новая специальность “Электрооборудование промышленных предприятий” (прием 2 группы).

Ни в коей мере не умаляя заслуг преподавателей, инженерных работников в организации и становлении факультета в первые 10-15 лет, когда речь в основном шла о подготовке инженерных кадров и решении сложнейших конструкторских и производственных задач, отметим, что научные исследования на факультете начали интенсивно развиваться в пятидесятых годах и инициатором, организатором и руководителем их был выдающийся ученый-радиотехник Д.В. Агеев, приглашенный в индустриальный институт в 1949г.

Как подлинно выдающийся ученый Дмитрий Васильевич буквально в нескольких работах на десятилетия определял перспективные научные направления и открывал широкие горизонты своим многочисленным ученикам. Им была разработана теория групповой передачи сигналов с использованием их разделения по форме, позволившая на 1-2 порядка повысить помехоустойчивость радиотехнических системы и заложить основу для создания систем радиосвязи OFDM. Этот принцип в наши дни нашёл применение, реализован и закреплён в виде перспективного энергосберегающего и экологически чистого стандарта мобильной связи CDMA.

Д.В. Агеевым была разработана теория слеящего приема, предложен высокоэффективный ключевой метод усиления сигналов, выполнены классические исследования по спектральному анализу сигналов и радиотехнических систем, фазовой телеграфии.

Своими научными результатами Д.В. Агеев на много лет опередил американских ученых (К. Шеннона, Л. Заде и др.)

Д.В. Агеев внес неоценимый вклад в подготовку инженерных и научных работников, возглавляя в 1955-59 г.г. факультет, подготовив 14 докторов и более 50 кандидатов наук. Дмитрию Васильевичу было присвоено почетное звание Заслуженного деятеля науки и техники РСФСР, он был награжден орденом Трудового Красного Знамени.

Широкую известность получили научные результаты учеников Д.В. Агеева – профессоров Ю.С. Лезина, А.А. Горбачева, Ю.Н. Бабанова, Я.Г. Родионова, Ю.И. Пахомова

Юрия Сергеевича Лезина по праву следует считать основоположником и руководителем научных направлений в г. Горьком по цифровой обработке радиолокационной информации, электромагнитной совместимости РЭС (Полов К.П., Лещинский М.М., Курепин В.П., Горюнов М.В., Баранов В.Г., Кротов И.Д., Пантелеев Ю.В., Рындык А.Г.). В 1972-1989 г.г. Ю.С. Лезин – ректор политехнического института.

В развитии и становлении радиофакультета большое, возможно, определяющее значение сыграла кафедра радиотехники, которой заведовали Л.В. Львович (до 1937 г.), И.С. Белый (1938-1947 г.г.), В.Я. Сморгонский (1947-1949г.г.), Д.В. Агеев (1949-1952 г.г.), Г.В. Глебович (1952-1983 г.г.) – вместе с Л.А. Моругиным ставший основоположником разработок наносекундной импульсной техники, В.В. Крылов (1983-2004 г.г.), В.И. Есипенко (2004 -). На кафедре работал и получил звание профессора академик РАН А.В. Гапонов-Грехов.

В 1952 году радиофакультет выделился из электротехнического. Были созданы новые кафедры “Радиоприемные устройства” (заведующий – профессор Д.В. Агеев), “Радиопередающие устройства” (заведующий – Л.Н. Осташкин).

В 1959г. факультет возглавил доцент Л.Н. Осташкин. Леонид Николаевич Осташкин внес большой вклад в становление и развитие факультета. Выпускник (1930г.) Нижегородского механико-машиностроительного института по специальности “Электросиловые установки” ярко проявил себя в ЦВИРЛ, НИИ-11, получив дополнительно глубокие профессиональные знания и навыки организаторской работы. С 1946г. в Горьковском индустриальном институте (на спецфаке). Руководство института доверяло Л.Н. Осташкину организацию и заведование тремя кафедрами факультета и в течение почти 14 лет руководство ими. При Л.Н. Осташкине начали формироваться тесные научные и производственные связи между факультетом и отраслевыми НИИ и КБ.

В 1964г. под руководством Л.Н. Осташкина была создана новая кафедра “Промышленная электроника”, которой он руководил до 1972г. Тогда на факультете работали 5 профессоров, докторов и около 30 кандидатов наук. За большой вклад в развитие высшего образования Леонид Николаевич был награжден орденом “Знак Почета”. В 1977 году Л.Н. Осташкин вышел на пенсию.

В (50 ÷ 70)-х годах на факультете была организована подготовка инженеров по конструированию и производству радиоаппаратуры (1956 г.), математическим и счетно-решающим приборам и устройствам, автоматизированным системам управления, в связи с чем были образованы соответствующие кафедры. В 90-х годах были образованы кафедры “Системы передачи и обработки информации в сетях ЭВМ” (ныне кафедра “Электроника и сети ЭВМ”), “Графические информационные системы”, “Техника СВЧ”, “Теория цепей и телекоммуникаций”. Большой вклад в становление научно-педагогических коллективов внесли заведовавшие кафедрами профессора Ю.С. Бажанов, В.Г. Баранов, Ю.Г. Белов, В.Ф. Благов, Ю.К. Богатырев, В.И. Есипенко, Д.А. Кабанов, В.А. Калмык, В.В. Кондратьев, В.В. Крылов, И.Н. Мерзляков, С.Н. Митяков, В.Р. Милов, Л.А. Моругин, С.М. Никулин, Н.С. Петрухин, С.Б. Раевский, А.Г. Рындык, В.И. Сагунов, С.Г. Сапфилов, Р.М. Сидорук, Э.С. Соколова, А.И. Туркин.

В 2005 году на базе факультета образован учебно-научный институт радиоэлектроники и информационных технологий.

Основания для преобразования факультета в институт были весомые:

1. Значительный контингент студентов (2,5 тыс.) и широкий спектр направлений подготовки специалистов (14).
2. Наличие высококвалифицированного профессорско-преподавательского состава (40 докторов и 130 кандидатов наук).

3. Наличие научных школ, успешное проведение исследований по приоритетным научным направлениям и эффективная подготовка и аттестация научных кадров в 4-х докторских советах института.

4. Установившееся стратегическое партнерство между факультетом и федеральными научно-производственными центрами (НИИИС им. Ю.Е. Седакова, ОАО “НИИРТ”, “НПП “Полет”, НИПИ “Кварц им. А.П. Горшкова”), а также крупными телекоммуникационными компаниями (МЕРА, ТЕЛЕКА, ТЕКОМ).

Перечисленные обстоятельства позволяли факультету, а теперь институту успешно решать задачу подготовки инженерных и научных кадров на основе органичного соединения учебного процесса и научно-исследовательских и опытно-конструкторских работ.

С 2011 года в институте реализуется подготовка бакалавров, магистров и дипломированных инженеров по 6-ти направлениям, 14 профилям и специальности “Радиоэлектронные системы и комплексы”. В 2020 г. получена лицензия на открытие подготовки по программе специалитета по направлению «Информационная безопасность автоматизированных систем» и открыта новая кафедра ИРИТ: «Информационная безопасность вычислительных систем и сетей».

На сегодняшний день в структуре ИРИТ 8 кафедр, обеспечивающих подготовку около 2-х тыс. студентов, 300 магистрантов и 95 аспирантов:

- кафедра «Информационные радиосистемы» (заведующий – д.т.н., проф., заслуженный конструктор РФ А.Г. Рындык) направление “Радиотехника”, специальность “Радиоэлектронные системы и комплексы”;
- кафедра «Компьютерные технологии в проектировании и производстве» (заведующий – д.т.н., профессор С.Л. Моругин), направления “Проектирование и технология радиоэлектронных средств”, “Информационные системы и технологии”;
- кафедра «Вычислительные системы и технологии» (заведующий – д.т.н., доцент Д.В. Жевнерчук), направление подготовки “Информатика и вычислительная техника”;
- кафедра «Информатика и системы управления» (заведующий – д.т.н., проф. Э.С. Соколова), направления подготовки “Информатика и вычислительная техника”, “Информационные системы и технологии”;
- кафедра «Информационная безопасность вычислительных систем и сетей» (и.о. заведующего – к.т.н., доцент Ляхманов Д.А.) специальность «Информационная безопасность автоматизированных систем».
- кафедра «Прикладная математика» (заведующий – д.ф.-м.н., проф. А.А. Куркин) направление “Прикладная математика и информатика”;
- кафедра «Электроника и сети ЭВМ» (заведующий – д.т.н., проф. В.Р. Милов), направления “Информационные системы и технологии”, “Инфокоммуникационные системы и сети связи”;
- кафедра «Графические информационные системы» (заведующий – к.т.н., доцент А.Д. Филинских), направление “Информационные системы и технологии”

Научные школы и направления научных исследований

На факультете на интуитивном уровне в конце 40-х – 50-х годов, когда создавался ракетно-ядерный щит СССР, самостоятельно формировалась система подготовки кадров, сущность которой заключается в органичном соединении учебного процесса, научных исследований и конструкторских работ. Это было обусловлено тем, что факультет всегда являлся и является сейчас базовым подразделением высшего образования для оборонно-промышленного комплекса, требующего высококвалифицированных специалистов.

Проектирование высокоэффективных радиолокационных систем и комплексов

Первую научную школу факультета создал и руководил ею Д.В. Агеев (см. выше). Эта школа по помехоустойчивости радиотехнических систем и комплексов различного назначения получила широкое признание в СССР, за рубежом. Практически все разработки были оригинальными защищены авторскими свидетельствами на изобретения и внедрены в серийно выпускавшихся изделиях. В настоящее время на кафедре информационных радиосистем под руководством профессора А.Г. Рындыка и при активном участии профессоров А.Д. Плужникова, А.В. Мякиньюкова, А.В. Андриянова, А.Б. Бляхмана, А.Г. Флаксмана проводятся интенсивные НИОКР по созданию систем обработки информации средствами ПВО на отечественной элементной базе. Разрабатывается новый метод радиолокации “на просвет”.

Докторские диссертации в рамках указанного направления защитили Ю.И. Пахомов, А.Г. Рындык, А.Д. Плужников, А.Б. Бляхман, А.В. Мякиньюков, Ф.Н. Ковалев.

На базе кафедры образован университетский “Центр цифровых технологий” (руководитель – зав. кафедрой, профессор А.Г. Рындык), нацеленный на реализацию разработок в интересах ОПК. Кафедра и центр неоднократно выигрывали объявляемые Правительством конкурсы на проведение НИР и НИОКР. В настоящее время исследования и разработки ЦЦТ ориентируются на создание отечественных авторадаров для применения, в том числе, в беспилотных транспортных средствах.

Нанoeлектроника и измерительная техника

В научном, историческом плане трудно переоценить исследования, проведенные на кафедре радиотехники. У истоков научных исследований были профессора Г.В. Глебович, Л.А. Моругин, Ю.С. Лезин, академик А.В. Гапонов-Грехов.

Единое научное направление кафедры стало формироваться с 1955г. К этому времени доценты Г.В. Глебович и Л.А. Моругин выполнили фундаментальные исследования по разработке теории и устройств наносекундной импульсной техники, остро востребованных радиоизмерительной отраслью, ими опубликованы монографии «Формирование импульсов наносекундной длительности» (1958), и «Наносекундная импульсная техника» (1964), ставшие настольными книгами многих радиоинженеров.

В 70-х годах сложились и развились исследования в области применения наносекундной и пикосекундной техники в радиоизмерительной технике – импульсной рефлектометрии, начаты разработки информационно-измерительных систем.

За цикл работ по автоматизированным системам проведения экспериментов во временной области группа сотрудников кафедры В.В. Крылов (руководитель), А.В. Андриянов, В.Д. Глушков, Е.Б. Марамчина, С.Я. Корсаков, Д.М. Пономарев удостоены в 1979 г. Премии Ленинского комсомола.

Докторские диссертации защитили И.Г. Катаев, Д.А. Кабанов, Ю.К. Богатырев, В.В. Крылов, М.И. Грязнов, Ю.А. Рябинин.

С 1983г. под руководством д.ф.-м.н., профессора В.И. Есипенко проводятся исследования по статистической теории динамических систем.

После перехода Л.А. Моругина на кафедру «Конструирование и производство радиоаппаратуры» (ныне «Компьютерных технологий в проектировании и производстве») к исследованиям начали привлекаться специалисты из промышленности.

Под научным руководством Л.А. Моругина подготовил докторскую диссертацию Ю.Е. Седаков впоследствии Заслуженный деятель науки и техники РФ, лауреат Государственной премии, директор НИИ измерительных систем. Ю.Е Седаков руководил аспирантурой на кафедре КИТР в области систем автоматизированного проектирования и автоматизации технологической подготовки производства.

Естественным развитием научных интересов Л.А. Моругина и Д.А. Кабанова стали исследования волновых процессов в линейных, нелинейных и параметрических электромагнитных системах с распределенными параметрами. Результатом исследования

стала Монография Д.А. Кабанова «Функциональные устройства с распределенными параметрами» (издательство «Советское радио», 1979). На кафедре появились новые научные лидеры Е.А. Егоров, С.М. Никулин, С.Л. Моругин, которые вместе с Д.А. Кабановым организовали исследовательскую работу молодых сотрудников: С.М. Никулин развивает теорию и методы автоматизированных измерений волновых параметров интегральных СВЧ-структур; Моругин С.Л. включается в разработку резистивных пленочных элементов микроэлектроники и новых методов анализа нелинейных элементов и устройств ВЧ- и СВЧ-диапазона.

Докторские диссертации защитили С.М. Никулин, С.Л. Моругин, В.П. Хранилов, А.В. Баранов, М.А. Ивлев.

При кафедре совместно с Арзамасским приборостроительным заводом создана НИЛ «Автоматизированное проектирование радиоэлектронных систем и микроволновые измерения».

Системы управления и обработки информации

Ведущей научной школой не только ИРИТ, университета и Н. Новгорода является школа, созданная чл.-корр. РАН, зав. кафедрой вычислительных систем и технологий В.В. Кондратьевым в области теории и методов проектирования цифровых систем управления и обработки информации. Результаты исследований школы позволили создать ряд систем по управлению сложными объектами и технологическими комплексами и реализовать их на предприятиях Н. Новгорода, Ижевска, Воронежа, Подмосквья.

Огромен вклад научной школы в подготовку научных кадров. При научном консультировании и руководстве В.В. Кондратьева защищено свыше 30 докторских и около 70 кандидатских диссертаций.

Под руководством профессора В.А. Утробина проводятся исследования по проблеме построения информационных модулей, изоморфных на врожденные механизмы системы зрительного восприятия и позволяющих формализовать процедуры композиционного анализа изображений.

Исследованием методов верификации и тестирования программных систем, построения моделей тестовых данных и создания на их основе систем обработки и анализа текстов успешно занимается группа ученых под руководством профессора Л.С. Ломакиной.

Докторские диссертации защитили В.А. Утробин, Ю.С. Бажанов, Г.В. Кондратьев, П.В. Мисевич, А.С. Суркова, Д.В. Жевнерчук.

На кафедре длительное время проводились исследования по электромагнитной совместимости РЭС и диагностике цифровых схем (В.Г. Баранов, А.Н. Александров, Ю.С. Бажанов, П.И. Уваров), а также по АЦП и ЦАП (В.И. Александрин, Ю.Н. Киселев).

Техническая диагностика непрерывных и дискретных объектов

Научная школа «Техническая диагностика непрерывных и дискретных объектов» сформировалась на кафедре «Информатики и систем управления» под руководством д.т.н., проф., заслуженного деятеля науки РФ В.И. Сагунова. Исследования, включают три основных направления: синтез контролепригодных объектов непрерывного действия путем определения оптимальных совокупностей диагностических параметров, разработка автоматизированных процедур контроля работоспособности и поиска дефектов произвольной кратности; разработка и исследования методов, процедур и алгоритмов построения тестов для проверки и диагностирования цифровых схем; разработка методов диагностирования программных средств в системах обработки информации и управления.

Под руководством В.И. Сагунова были выполнены крупные бюджетные и хоздоговорные НИР, заказчиками которых являлись многие крупные предприятия г. Н. Новгорода, Москвы и Московской области.

У истоков создания школы В.И. Сагунова (период с 1972 г. по 1985 г.) стояли его ученики - С.И. Гаркавенко-Беляева, В.С. Нечаев, Л.С. Ломакина, А.Е. Миндров.

Докторские диссертации защитили С.И. Беляева, Л.С. Ломакина, Э.С. Соколова.

Исследования, успешно продолжаются под руководством нынешней заведующей кафедрой, профессора Э.С. Соколовой и, в свете современных требований к информационным системам, ориентируются на решение проблем безопасности информации на основе методов биометрии и распознавания изображений, систем технического зрения и искусственного интеллекта, а также эффективной обработки сверхбольших объемов данных.

Прикладная и теоретическая электродинамика

Формирование школы прикладной электродинамики в ГПИ им. А.А. Жданова относится к 1960-1970 г.г., когда заведующий кафедрой проф. В.Я. Сморгонский начал подготовку аспирантов по технике сверхвысоких частот. С приходом в аспирантуру талантливых молодых специалистов С.Б. Раевского, Ю.А. Иларионова, а затем Ю.Г. Белова, В.А. Калмыка, А.А. Радионова, Е.П. Тимофеева и др. исследования по прикладной электродинамике и технике СВЧ развернулись широким фронтом.

Значительным вкладом в теорию гофрированных и частично заполненных диэлектриком круглых и эллиптических волноводов стала книга, написанная В.Я. Сморгонским, С.Б. Раевским и Ю.А. Иларионовым, «Расчет гофрированных и частично заполненных волноводов», которая и поныне является рабочим пособием инженеров, проектирующих устройства СВЧ.

Под научным руководством Заслуженного деятеля науки РФ, профессора, д.т.н., С.Б. Раевского защитили докторские диссертации А.А. Радионов, В.А. Калмык, Ю.Г. Белов, А.С. Раевский, В.А. Козлов, А.В. Кашин, А.А. Титаренко, А.Ю. Седаков, В.В. Бирюков, В.А. Малахов. Под руководством проф. А.А. Радионова защитил докторскую диссертацию Р.В. Бударагин.

Школе принадлежат: разработка и создание теории комплексных волн в направляющих электродинамических структурах, определение общих подходов к исследованию электродинамических структур, описываемых несамосопряженными операторами, создание методов исследования неоднородных и нерегулярных направляющих структур (в том числе с резистивными и сверхпроводящими пленками), невзаимных направляющих структур, алгоритмов расчета и параметрического синтеза волоконных световодов, в том числе периодически-нерегулярных; предложен и развит целый ряд методов прикладной электродинамики: МЧО с непрерывным спектром собственных функций, метод интегральных уравнений на основе леммы Лоренца, метод поверхностного тока, модифицированный метод Галеркина, метод вариации фазы, корреляционный метод коллокаций. Открыто и определено явление комплексного резонанса. Разработанные методы позволяют производить электродинамический расчет в принципе произвольных направляющих структур, в частности, линий передачи терагерцового диапазона. Издано более 10 монографий и 5 учебников.

Интеллектуальные информационные системы

На рубеже веков (1998-2005г.г.) на кафедре “Электроника и сети ЭВМ” под руководством чл.-корр. РАН В.В. Кондратьева, д.т.н., проф. В.Р. Милова, к.т.н., проф. В.Г. Баранова произошло становление нового научного направления «Интеллектуальные и нейросетевые системы обработки информации и принятия решений», включающего нейроинформатику и интеллектуальный анализ данных, распознавание образов и поддержку принятия решений при наличии существенной неопределенности, моделирование и интеллектуализацию систем и процессов управления, разработку алгоритмического обеспечения и программную реализацию интеллектуальных систем.

В 2012 г. на базе кафедры (заведующий – проф. В.Р. Милов) создан университетский научно-технический инновационный центр «Радиоэлектронные системы и информационные технологии в промышленности».

Выполнены исследования в области моделирования телекоммуникационных сетей, разработки алгоритмического и программного обеспечения для систем радиосвязи с подвижными объектами. Разработано программное обеспечение расчета и визуализации зон радиопокрытия базовых станций

Проводятся исследования и разработки информационно-аналитических систем, алгоритмов и технологий мониторинга распределенных технических объектов. В рамках работ по этому направлению в 2012-14 г.г. проведен комплекс работ по разработке подсистемы обработки данных и автоматизированного рабочего места специалиста по геотехническому мониторингу интеллектуальной информационной системы комплексного мониторинга состояния магистрального газопровода (ИИС КМСГ). Успешно завершены пуско-наладочные работы на объекте «Магистральный газопровод «Сахалин-Хабаровск-Владивосток», в процессе которых на региональном диспетчерском пункте Хабаровского ЛПУМГ выполнена настройка программных компонентов ИИС КМСГ и их сопряжение с подсистемами источниками данных.

В 2014-2016 г.г. под редакцией к.т.н. В.Г. Баранова и д.т.н. В.Р. Милова в издательстве «Радиотехника» (г. Москва) изданы 3 тома серии научных монографий «Разработка и моделирование информационных систем».

Прогнозирование природных катастроф

На кафедре прикладной математики создана и функционирует научная школа по методам прогнозирования природных катастроф под руководством д.ф.-м.н. профессора А.А. Куркина. В работе, организованной при кафедре научно-исследовательской лаборатории моделирования природных и техногенных катастроф активное участие принимают специалисты Института прикладной физики РАН

Объектами исследований, получивших признание в стране и за рубежом, являются:

- Волны цунами, их математическое описание и прогнозирование характеристик волн, необходимое для развития хозяйственной деятельности на побережьях океанов и морей;
- Волны-убийцы на поверхности моря, возникающие неожиданно на короткое время зачастую в спокойном море;
- Краевые волны, распространяющиеся вдоль берега и оказывающие влияние на изменение линии берега. Впервые показано, что в краевых волнах также возможно образование аномальных волн-убийц, приводящих к неожиданным заплескам волн на берег;

Разработана модель динамики поверхностных загрязнений в поле волн и течений, позволяющая предсказать их распространение, а также служить индикатором крупномасштабных волновых процессов.

Докторские диссертации защитили А.А. Куркин, Р.Х. Мазова, Л.Ю. Катаева, И.И. Диденкулова, А.С. Козелков, А.И. Зайцев, А.В. Радостин.

Руководитель научных исследований профессор А.А. Куркин удостоен многочисленных наград и премий, в т.ч. медали Плиниуса Европейского геофизического союза. Молодой доктор наук И.И. Диденкулова в 2015 г. была удостоена премии Президента РФ для молодых ученых в области науки и инноваций. И.И. Диденкуловой и профессору Е.Н. Пелиновскому также были присуждены медали Плиниуса. Коллектив успешно участвует в международных и российских выставках и конкурсах, занимая призовые места и завоёвывая дипломы победителей.

Учёные ИРИТ регулярно становятся победителями конкурсов грантов различных научных фондов. Так грантами Президента были удостоены И.И. Диденкулова (2012-2013 г.г., 2014-2015 г.г., 2016-2017 г.г.), А.А. Куркин (2016-2017 г.г., 2020-2021 г.г., 2020-2021 г.г.), А.Р. Гиниятуллин (2016-2017 г.г., 2018-2019 г.г., 2020-2021 г.г.), А.И. Зайцев (2020-2021 г.г.), А.С. Козелков (2018-2019 г.г.).

Гранты РФФИ выигрывали А.А. Куркин (2012, 2014-2016 г.г.), И.И. Диденкулова (2012-2013 г.г., 2015-2016 г.г. – 2 гранта, А.В. Мякинков (2015-2016 г.г.), Е.А. Рувинская (2016-2017 г.г.), (А.С. Козелков 2016-2017 г.г.), В.П. Хранилов (2014, 2015, 2017, 2019-2021 г.г., 2020 г.), Г.Б. Бронфельд, Д.И. Киров (2020-2021 г.г.).

Грант РФФИ И.И. Диденкулова (2016-2018 г.г.).

Получали финансирование ФЦП Минобрнауки РФ: Куркин А.А. (2010-2012 г.г., 2011-2012 г.г., 2014-2016 г.г.), И.Н. Малышев (2010-2012 г.г.), В.В. Крылов, Э.С. Соколова (2014-2015 г.г.), кафедры ИСУ и ЭСВМ (2015-2017 г.г.), кафедра ЭСВМ (2016-2017 г.г.).

Финансирование по госпрограмме РФ «Развитие науки и технологий на 2013-2020 г.г.» (пост. №218): кафедра ИРС А.Г. Рындык, А.В. Мякинков (2012-2015 г.г., 2016-2018 г.г., 2019-2021 г.г.).

Для института характерна активная издательская деятельность: В 1995-2010 г.г. выпускаемые в ИРИТ тематические научные сборники «Радиоэлектронные и телекоммуникационные системы и устройства», «Системы обработки информации и управления», «Информационные технологии» фактически формировали текущие тома научно-технического журнала «Труды НГТУ», включенного впоследствии в список изданий, рекомендованных ВАК, для публикации результатов диссертаций. Только за последние 5 лет издано более 30 монографий, опубликовано свыше 200 научных статей в журналах с высоким рейтингом, индексируемым в международных системах цитирования.

Институт регулярно участвует в издании “журналов в журналах” (“Информационно-измерительные и управляющие системы”, “Нейрокомпьютеры: разработка и применение”, “Системы управления и информационные технологии”, “Датчики и системы”, “Антенны”).

В 4-х диссертационных советах, созданных на базе института (председатели В.В. Кондратьев, А.С. Раевский, А.Г. Рындык, Н.С. Петрухин) активно проводятся защиты кандидатских и докторских диссертаций научно-педагогическими работниками института, НГТУ, других вузов, включая ведущие столичные, а также сотрудниками федеральных исследовательских центров и крупных предприятий. Основной состав в 2-х объединённых диссертационных советах, созданных НГТУ им. Р.Е. Алексеева совместно с “НПП “Полет” (зам. председателя от НГТУ В.Р. Милов) и ННГАСУ (зам. председателя от НГТУ В.П. Хранилов) составляют профессора ИРИТ.

Институт ежегодно проводит Международную научно-техническую конференцию «Информационные системы и технологии», в которой участвуют исследователи не только из РФ, но и стран дальнего и ближнего зарубежья (конференция поддерживается грантами РФФИ). Большое внимание на конференции уделяется апробации научных работ молодых ученых. На сегодняшний день проведено 27 конференций.

За 26 лет существования конференции ИСТ докторские диссертации защитили более 30 участников конференции разных лет. Среди таких профессора кафедр ИРИТ, других факультетов и институтов НГТУ и вузов Нижнего Новгорода: С.Л. Моругин, В.А. Утробин, Г.В. Кондратьев, Э.С. Соколова, Ю.С. Бажанов, С.А. Лабутин, А.А. Куркин, В.В. Андреев, Л.Ю. Катаева, В.П. Хранилов, П.В. Мисевич, А.В. Мякинков, А.В. Баранов, О.А. Морозов, М.А. Ивлев, И.И. Диденкулова, Ф.Н. Ковалев, А.С. Суркова, Ю.М. Туляков, А.В. Савченко, Н.Ю. Бабанов, А.В. Слюняев, А.В. Синелева, Д.В. Жевнерчук, В.В. Поздьяев, А.С. Козелков, А.И. Зайцев, А.В. Радостин, М.В. Ширяев, Н.В. Старостин, К.В. Баркалов.

В основном это доктора технических наук, но есть и 10 докторов физико-математических наук, доктор исторических наук (Зубков А.А.), доктор филологических наук (Синелева А.В.), доктор экономических наук (М.В. Ширяев). Причем 14 из нынешних докторов (Г.В. Кондратьев, В.Р. Милов, В.В. Андреев, А.С. Раевский, П.В. Мисевич, А.В. Мякинков, А.В. Баранов, Ф.Н. Ковалев, А.В. Савченко, А.В. Синелева,

А.С. Суркова, В.В. Поздьяев, А.С. Козелков, А.И. Зайцев) за прошедшие 26 лет прошли путь от аспиранта до доктора наук.

Кандидатские диссертации защитили более 100 человек. Из них в диссертационных советах НГТУ более 80. Здесь, в основном, кандидаты технических и физико-математических наук, но есть и 2 канд. медицинских наук (Пестрикова В.Н., Колчина Н.С.), кандидат педагогических наук (Баранова А.В.), кандидат и доктор филологических наук (Синелева А.В.) кандидаты химических наук, и даже кандидат технических наук, защитившая ученую степень кандидата искусствоведения (Садкова О.В.).

Научные коллективы, работники, аспиранты, магистранты и студенты активно участвуют в проводимых Правительством РФ и Минобрком конкурсах, являются их победителями, стипендиатами Президента и Правительства. Новые подходы создают новые условия для развития интереса и приобретения высокотехнологичных компетенций в ИТ. Студенты ИРИТ неоднократно становились победителями Всероссийского чемпионата CASE-IN в номинации “Цифровой атом”, Всероссийского конкурса «Цифровой прорыв» для профессионалов ИТ-отрасли всех уровней, Всероссийских хакатонов VirusHack от Ростелекома, TenderHack в номинации поиск по фото, «Интеллектуальные транспортные системы и элементы ситуационных центров», Intel Machine Learning Optimization Challenge 2020, Хакатон Труда, «ИТ-проект года» 2017-2019 г.г. Команды кафедр ИСУ и ЭСВМ ИРИТ принимают активное участие в мероприятиях ежегодного инновационного конкурса «УМНИК» и областного конкурса «РОСТ», выигрывая гранты и призовые места.

Успехами в реализации эффективной системы подготовки научных и инженерных (теперь бакалавров, магистров) кадров совместно и в тесном сотрудничестве с промышленностью институт обязан тесному сотрудничеству со своими стратегическими партнерами: Федеральными научно-производственными центрами: “НИИИС им. Ю.Е. Седакова”, “НПП “Полет”, ОАО “НИИРТ”, НИИПИ “Кварц им. А.П. Горшкова”, рядом заводов, телекоммуникационных компаний, в которых организованы базовые кафедры, и филиалы кафедр ИРИТ. Первый в Горьковской области филиал кафедры КИТР (сегодня КТПП) в НПП «Полёт» был открыт в 1982 г. Далее были филиалы кафедр ИРИТ в НИИИС, НИИПИ, ГИ «Салют», НИИИРТ.

Институт гордится своими выпускниками, в числе которых 7 лауреатов Ленинской премии (В.В. Карпов, А.М. Клячев, К.И. Ложкарев, Н.В. Мохин, Ю.Н. Соколов, Б.Ф. Челышев, Г.С. Ушаков), более 50 лауреатов Государственных премий СССР, РФ, десятки докторов и сотни кандидатов наук, руководители крупных объединений, научных центров, предприятий. С.В. Гапонов избран академиком РАН, Н.Н. Салашенко – чл.-корр. РАН. Они же удостоены Государственных премий СССР.

За успехи в научно-педагогической деятельности Государственных наград удостоены сотрудники факультета, института: Д.В. Агеев – орден “Трудового Красного знамени”; В.Г. Баранов – ордена “Знак Почета”, “Почета”; Г.В. Глебович – орден “Знак Почета”; Ю.С. Лезин – ордена “Октябрьской революции”, “Трудового Красного знамени”; В.В. Маланов – орден “Знак Почета”, медаль “За трудовую доблесть”; Л.Н. Осташкин – орден “Знак Почета”; В.Я. Сморгонский – медаль “За доблестный труд”.

Почетных званий удостоены: “Заслуженный деятель науки и техники РФ” – Д.В. Агеев, Д.А. Кабанов, В.В. Кондратьев, Ю.С. Лезин, В.И. Сагунов; “Заслуженный конструктор РФ” – А.Г. Рындык; “Заслуженный работник высшей школы РФ” – В.Г. Баранов, С.М. Никулин; “Почётный работник высшего профессионального образования” – В.Г. Баранов, Н.И. Великанов, В.И. Есипенко, В.В. Кондратьев, Л.С. Ломакина, И.Н. Мерзляков, А.Д. Плужников, Н.Е. Пособилов, А.В. Семашко, В.П. Хранилов. “Почётный работник сферы образования РФ” – М.А. Ивлев, В.Р. Милов, Э.С. Соколова; «Почётный радист» – Д.В. Агеев, В.Г. Баранов, А.И. Гречихин, В.И. Есипенко, А.В. Семашко, В.П. Хранилов, А.М. Шабалин.

Звания “Лауреат премии Ленинского комсомола” удостоены В.В. Крылов, А.В. Андриянов, В.Д. Глушков, Е.Б. Марамчина, С.Я. Корсаков, Д.М. Пономарев.

Звания “Лауреат премии г. Нижнего Новгорода” удостоены В.Г. Баранов, Б.Д. Забегалов, Н.С. Новиков, Э.М. Ямпольский.

Накопленный интеллектуальный и производственный опыт, активная позиция научно-педагогических и инженерных работников института вселяют уверенность в его успешном развитии и позиционировании в будущем в составе Нижегородского государственного технического университета им. Р.Е. Алексеева (Опорного университета).

Декан факультета с 1975г., директор ИРИТ (до 2017 г.), Заслуженный работник высшей школы РФ, Почётный работник высшего профессионального образования РФ, Почетный Радист РФ, к.т.н., профессор В.Г. Баранов

Зам. декана факультета с 1990 г., зам. директора ИРИТ с 2007г., Почётный работник высшего профессионального образования РФ, Почетный Радист РФ, д.т.н., профессор В.П. Хранилов

Библиографический список

1. Баранов В.Г. 75 лет ИРИТ. Из славного прошлого в светлое и успешное будущее / В.Г. Баранов // Материалы XVII международной науч.-техн. конф. «Информационные системы и технологии», ИСТ-2011, Н. Новгород, НГТУ, 2011, С. 5-14.
2. Баранов В.Г. Институт радиоэлектроники и информационных технологий / В.Г. Баранов // Материалы XXII международной науч.-техн. конф. «Информационные системы и технологии», ИСТ-2016, Н. Новгород, НГТУ, 2016, С. 5-12.
3. Баранов В.Г. 100 лет Нижегородскому политехническому институту. Радиофакультет в Политехе. / В.Г. Баранов // Материалы XXIII международной науч.-техн. конф. «Информационные системы и технологии», ИСТ-2017, Н. Новгород, НГТУ, 2017, С. 5-10.
4. Баранов В.Г. Радиофак на защите отечества / В.Г. Баранов, В.П. Хранилов, // Материалы XXVI международной науч.-техн. конф. «Информационные системы и технологии», ИСТ-2020, Н. Новгород, НГТУ, 2020, С. 4-9.

V.G. Baranov, V.P. Khranilov

85 YEARS OLD THE RADIO FACULTY IN POLITECH. HISTORY AND TODAY

Nizhny novgorod state technical university n.a. R.E. Alekseev

The history of the creation and development of the radio engineering faculty in the Nizhny Novgorod Polytechnic Institute (ISTU named after RE Alekseev) is presented. The structure and composition of the Institute of Radioelectronics and Information Technologies are described. The information on the leading scientific and pedagogical staff is presented. Information about the awards of outstanding employees and graduates is given.

Keywords: Nizhny Novgorod Polytechnic Institute, Radio Engineering Department, Gorky Industrial Institute, Faculty of Radioelectronics and Technical Cybernetics, Faculty of Information Systems and Technologies, Institute of Radio Electronics and Information Technologies.

**ЮБИЛЕЙ ВЕЛИКОГО УЧЕНОГО И ПЕДАГОГА ПРОФЕССОРА
ДМИТРИЯ ВАСИЛЬЕВИЧА АГЕЕВА**

Нижегородский государственный технический университет им. Р.Е. Алексеева

*...может собственных Платонов
И быстрых разумом Невтонов
Российская земля рождать.*

М.В. Ломоносов

Профессор Д.В. Агеев является основоположником теории связи и теории передачи информации, автором более 10 новых научных направлений в теоретической радиотехнике. Все результаты выполненных им теоретических исследований были пионерскими, то есть полученными впервые в мире. Под его научным руководством защищено 55 кандидатских диссертаций, 14 его учеников, в том числе автор данной статьи в дальнейшем стали докторами наук. В Горьком сформировалась радиотехническая научная школа Д.В. Агеева. Впервые в мире он разработал теорию и активный метод обучения студентов технических вузов. Из 55 учеников Д.В. Агеева 32, в том числе семь докторов наук были преподавателями. Фактически он создал факультет и вывел его на уровень старейших радиотехнических факультетов страны.

Ключевые слова: основоположник теории связи, научная школа, теория и метод обучения, создание факультета.

Введение

В связи с недавним столетием Нижегородского государственного технического университета были особо отмечены люди, составляющие его гордость и славу, добившиеся в своей деятельности в области науки и техники выдающихся результатов мирового уровня. Это создатель принципиально новых видов транспортной техники Р.Е. Алексеев, имя которого носит университет, создатель атомной энергетики и физико-технического факультета И.И. Африкантов, создатель новой техники полярный исследователь А.Ф. Николаев, профессор Д.В. Агеев – основоположник теории связи и теории передачи информации, создатель новых направлений в теоретической радиотехнике, научной школы, теории и метода обучения студентов и радиотехнического факультета. 21 февраля 2021 года исполнилось 110 лет со дня рождения Д.В. Агеева.

Жизненный путь и научная деятельность

Дмитрий Васильевич Агеев родился в Санкт-Петербурге в семье рабочего-литейщика, родители были выходцами из крестьян Тверской губернии. В школьные годы он увлекся радио, посвящал все свободное от учебы время изготовлению радиоприемников. В 1929 году после окончания школы-девятилетки он пытался поступить на учебу в Ленинградский электротехнический институт (ЛЭТИ), но не был принят. У него не было характеристики об общественной работе от комсомольской организации. Год Д. Агеев занимался на государственных подготовительных курсах. Его успехи в учебе на курсах были такими, что он получил рекомендацию поступать на математический факультет университета. Однако желание изучать радиотехнику оставалось неизменным. В 1930 году он был принят на радиотехнический факультет только что открытого Ленинградского электротехнического института связи (ЛЭИС). Студента Д.В. Агеева в ЛЭИС обучали выдающиеся ученые-педагоги своего времени:

основоположник отечественной радиотехники член-корреспондент АН СССР М.А. Бонч-Бруевич, профессора А.А. Пистолькорс, В.В. Татаринов. Исключительные творческие способности, неукротимое стремление познать непознанное, целеустремленность и трудолюбие в полной мере проявились у студента Д. Агеева и сохранились на всю его жизнь. На третьем курсе в 1933 году по собственной инициативе и полностью самостоятельно он написал свою первую научную работу «Обобщение метода Ньютона вычисления корней уравнения». В 1934 году она была опубликована в научно-техническом сборнике ЛЭИС. Эта работа упоминалась в вышедшей в 1937 году книге «Успехи советской математики за 20 лет». В 1934 году Д. Агеев написал новую работу, а затем на семинаре в ЛЭТИ услышал доклад доцента (впоследствии академика) Ю.Б. Кобзарева на ту же тему. Их совместная статья «О переходных процессах в резонансном усилителе», опубликованная в 1935 году в Журнале технической физики, открыла новое направление – исследование переходных процессов.

Затем студент Д.В. Агеев начал работать в области помехоустойчивости радиоприема и получил в ней самые существенные результаты. Они были обобщены в выполненной под руководством М.А. Бонч-Бруевича дипломной работе «Методы борьбы с помехами при радиоприеме», признанной государственной комиссией выдающейся. По результатам этой работы в сборнике ЛЭИС №10 за 1935 год была опубликована большая статья «Основы теории линейной селекции». Впервые полученные в ней главные результаты таковы. Необходимым и достаточным условием разделения сигналов является их взаимная линейная независимость. Принципов разделения только три: частотный, фазовый (временной) и компенсационный (разделение по форме сигнала, в дальнейшем кодовое разделение). В те годы было известно и применялось только частотное разделение сигналов. Два других принципа впервые предложены и исследованы Д.В. Агеевым. Определены требования к сигналам при использовании всех принципов их разделения. Эта статья легла в основу первой части диссертации «Теория линейной селекции и проблема пропускной способности «эфира», написанной Д.В. Агеевым в 1938 году. Частично аналогичные результаты американские ученые Л.А. Заде и К.С. Миллер получили в 1952 году [1]. Во второй части диссертации впервые решена важнейшая проблема пропускной способности канала связи, определена предельная скорость передачи. В основу теории положена физически реализуемая модель сигнала конечной длительности. Впервые процесс передачи и приема сигналов рассматривался с учетом влияния помех. Впервые использовалось векторное представление сигналов и помех в многомерном функциональном пространстве. В дальнейшем такое представление стало общепринятым. В 1938 году понятие информация и ее мера еще не были определены. Тем не менее вариационная трактовка проблемы и использование соответствующего математического аппарата позволили получить решение в простой физически прозрачной форме и применить его в телеграфии и телефонии. Американские ученые К.Э. Шеннон и Н. Винер определили пропускную способность канала связи в 1948 году [2,3].

В феврале 1939 года данная диссертация была защищена в качестве кандидатской в совете ЛЭИС и рекомендована к защите в качестве докторской в правомочном на это совете. Молодой ученый счел неэтичным дважды защищать одну и ту же работу. В развитие ее второй части он разработал теорию групповой передачи сигналов и изложил ее в докторской диссертации «Новый метод многоканального телеграфирования», которую защитил в 1940 году в ЛЭТИ. Фундаментальные результаты теоретических исследований, выполненных Д.В. Агеевым в 1935-1940 годах, до последнего времени не были опубликованы. О них знал только узкий круг отечественных ученых, за рубежом они не были известны. Только в 2017 году в связи со столетием Нижегородского государственного технического университета кандидатская диссертация Д.В. Агеева была издана небольшим тиражом [4]. Указанные выше результаты, как оказалось, опережали на 40-50 лет технические возможности их полной реализации. Однако они четко указывали цели и направления развития техники.

История теории связи и теории передачи информации началась со статьи (1935 г.) и диссертации (1938 г.) Д.В. Агеева, в этом приоритет у него, у России.

В нашей стране, начиная с 1942 года, результаты теоретических исследований Д.В. Агеева были положены в основу создания на базе имевшейся тогда техники секретной системы мобильной связи для высших эшелонов власти. Более подробная информация об этой системе и сведения о Дмитрие Васильевиче Агееве (не полностью соответствующие действительности) содержатся в телевизионном фильме 2005 года канала Культура «Мобильный для Лубянки».

Первую военную зиму Дмитрий Васильевич с женой и дочкой прожил в блокадном Ленинграде. ЛЭИС, в котором он работал, готовил связистов для армии. В марте 1942 года институт был эвакуирован сначала в Кисловодск, а затем в Тбилиси. Доцент Д.В. Агеев был заведующим кафедрой высшей математики, преподавал и математику, и радиотехнику. Он не вернулся в Ленинград, так как больной дочери холодный и сырой климат был противопоказан. В 1945 году он был переведен в Одессу в электротехнический институт связи заведующим кафедрой радиоприемных устройств. В 1949 году Дмитрий Васильевич был вынужден покинуть институт связи. Как единственный доктор технических наук в нем он был назначен рецензентом представленной к защите диссертации инженера-механика по специальности ректора института. Тема диссертации (подшипники скольжения) была устаревшей, результаты – ошибочными. Публикаций не было, вразумительных ответов на вопросы рецензента тоже. Диссертация была отклонена. Дмитрия Васильевича обвинили в антисемитизме, условия его работы стали нетерпимыми.

В сентябре 1949 года доктор технических наук Д.В. Агеев выбрал в качестве места работы Горьковский индустриальный институт, сейчас Нижегородский государственный технический университет имени Р.Е. Алексеева. В нем прошли почти 40 лет его неустанной творческой научной и педагогической работы.

Круг научных интересов и научной деятельности профессора Агеева был необычайно широк. Он внес существенный вклад в ряд направлений теоретической радиотехники, создал и развил новые направления.

Над теорией помехоустойчивости радиоприема он продолжал работать всю жизнь. Была создана теория следящего приема ЧМ сигналов, используемого в радиорелейной и спутниковой связи. Были исследованы системы связи с линейно независимыми неортогональными сигналами, с сигналами с взаимно-однозначным соответствием. Разработана теория разделения сигналов линейными фильтрами с переменными параметрами, определены их предельные возможности. Предложены и исследованы несколько видов нелинейных фильтров.

Теория спектральных представлений сигналов содержит научное обоснование всех основных понятий спектрального анализа. Введено новое понятие – активная полоса частотного спектра сигналов, вошедшее в книгу академика А.А. Харкевича «Спектры и анализ» и в дальнейшем использованное рядом ученых. Решена задача определения ширины спектра сигналов без вычисления структуры спектра с использованием не энергетического, а любого другого важного для практики критерия. Охвачены все виды импульсов и радиосигналов, в том числе модулированных.

Теория резонанса получила следующее развитие. Разработана спектральная теория резонанса в линейных пассивных и активных системах с постоянными параметрами, открывшая возможности создания новых устройств. Исследование резонанса в линейных системах с переменными параметрами позволило определить параметры сигнала, вызывающего резонанс в так называемом следящем фильтре с идеальным внешним управлением. Введено новое понятие – динамический резонанс, резонанс в колебательном контуре с нелинейным активным сопротивлением и зависящей от амплитуды колебаний частотой. Исследована возможность использования динамического резонанса для селекции импульсных радиосигналов.

Общая теория усиления радиосигналов основана на представлении усилителя как параметрической системы, находящейся под действием усиливаемого входного сигнала. Введено понятие об идеальном управляемом активном элементе, что позволяет развить аналитические методы исследования количественных соотношений между различными величинами, характеризующими работу усилителя. Теория так называемых ключевых методов усиления послужила основой повышения энергетической эффективности усилительных устройств.

Теория приближенных методов определения искажений сигналов охватывает все временные методы определения вносимых линейными цепями искажений видеосигналов и радиосигналов, в том числе квазистационарный метод и метод мгновенной частоты. Строго доказана необходимость прямого учета не вносящей искажений линейной части фазовой характеристики цепей.

В теории аппроксимации функций предложены более эффективные по сравнению с общеизвестными способами аппроксимации всех видов используемых видеоимпульсов и радиосигналов. Предложена и исследована система ортогональных гармонических функций с некратными частотами, частным случаем которой является ряд Фурье.

В общей теории модуляции и детектирования проанализированы многочисленные методы, по классификации профессора Д.В. Агеева виды модуляции: с переносчиками и без них, узкополосные и широкополосные, линейные и нелинейные, дифференциальные и интегральные, многозначная модуляция. Создана теория огибающих, предложены амплитудно-фазовый и энергетические детекторы.

Исследованы понятия частота и амплитуда. Показана неоднозначность представления функции $U(t)$ в виде колебания с переменной амплитудой $A(t)$ и частотой $\omega(t)$. Получен вывод о том, что эти величины должны быть наиболее медленно меняющимися из всех возможных для соответствия интуитивным представлениям о них. Строго доказана ошибочность теории о единственности определения частоты и амплитуды с помощью преобразования Гильберта.

Дмитрий Васильевич Агеев – ученый в самом высоком смысле этого слова. Теоретические научные исследования были его главным делом, он всегда был переполнен новыми идеями и замыслами. Ежедневно все свое свободное от работы время, в том числе выходные и отпуск, он посвящал этому любимому делу. В нем плодотворно сочетались глубокое проникновение в физическое содержание исследуемого явления и прекрасное владение математическим аппаратом для его описания. Целеустремленность и высокая работоспособность, творческая одаренность и высокий интеллект позволяли Дмитрию Васильевичу выбирать интересующую его проблему и доводить ее решение до требуемого результата.

Дмитрий Васильевич был полностью лишен тщеславия. В отличие от почти всех ученых для него главным, приносящим удовлетворение был результат исследования, а не его публикация. Он не считал целесообразным тратить время на подготовку публикаций, оформление документов. Закончив одно исследование, он сразу брался за другое. Будучи предельно добропорядочным и скромным, он не допускал, чтобы работу по оформлению за него делал кто-то другой. В общении с людьми, будь то студент или коллега, никогда не было и тени превосходства, только доброжелательность. Понимая необходимость передачи своих колоссальных знаний и опыта другим, Дмитрий Васильевич выбрал для этого уникальный, насколько известно, не примененный ни одним из ученых способ. В течение 13 лет еженедельно или раз в две недели он читал лекции для аспирантов, преподавателей, радиоинженеров города, т.е. для всех желающих, в которых излагал новые полученные им результаты. Он разрешал слушателям использовать их в своих исследованиях. Это был мастер-класс самого высокого научного уровня, школа творчества, научного поиска, логики научного исследования, использования различного, в том числе нового, математического аппарата.

С каждой лекцией слушатели поднимались на ступеньку вверх по лестнице знаний. Они поражались и восхищались оригинальностью решения сложных задач, доступностью изложения, ясным физическим смыслом полученного результата. У них появлялось и крепло желание заниматься научной работой и, естественно, иметь научным руководителем Дмитрия Васильевича. Выпускникам, оставленным в институте для обучения в аспирантуре, и поступившим в нее соискателям-преподавателям он щедро раздавал свои новые идеи (темы диссертаций), указывая направления и пути их развития. В дальнейшем к Дмитрию Васильевичу устремился поток аспирантов-заочников и соискателей с заводов, научно-производственных объединений и промышленных НИИ города со своими задачами и проблемами, решать которые им помогал научный руководитель. За почти 40 лет неустанной творческой научной и педагогической деятельности профессора Д.В. Агеева в Горьковском политехническом институте под его научным руководством 55 человек защитили кандидатские диссертации, а 14 из них затем и докторские.

Теоретические исследования Дмитрия Васильевича в большом числе научных направлений в радиотехнике и связи проложили столбовую дорогу, по которой вслед за своим научным руководителем пошли 55 его учеников. Главное из направлений – радиолокация. В ней наиболее успешно работали Ю.И. Пахомов и А.А. Зачепицкий. Сегодня в университете исследования продолжаются под руководством А.Г. Рындыка и А.В. Мякинкова. В радиофизическом институте под руководством А.А. Горбачева проведены работы по нелинейной радиолокации. Импульсную технику развивал Л.А. Моругин и сотрудники возглавляемой им кафедры. В области радиоизмерений активно работали А.В. Зенькович и А.М. Вирский, вопросы теории спектрально-временного анализа сигналов и систем успешно решались в Горьковском госуниверситете под руководством Ю.Н. Бабанова. Теорию и технику усиления сигналов развивали В.В. Маланов, В.Н. Ногин, Н.Б. Петяшин и Н.Б. Догадин, авиационную связь – И.Б. Петяшин. В радиофизическом институте, институте радиотехники, «Полете», «Кварце» работало несколько учеников Д.В. Агеева, в том числе доктора наук. В итоге к 1970 году сформировалась и далее быстро развивалась горьковская научная школа профессора Д.В. Агеева в области радиотехники.

Люди, знавшие работы Д.В. Агеева, и ученики считали его радиогением. В науке и технике к гениям относят ученых, имеющих достижения самого высокого уровня, выходящие за рамки общепринятых представлений, кажущиеся совершенно невозможными, парадоксальными. Писатель С. Цвейг утверждал: «Прекрасна истина, кажущаяся неправдоподобной». Такие истины представил миру Р.Е. Алексеев. А.С. Пушкин написал об этом так.

О, сколько нам открытий чудных
Готовят просвещенья дух
И опыт, сын ошибок трудных
И гений, парадоксов друг...

Профессор Д.В. Агеев доказал парадоксальную теорему, которую ученые поняли и приняли в книгах через 15-20 лет после доклада о ней на научной сессии в Москве. Ученый и историограф радиотехники и связи д.т.н. профессор М.А. Быховский говорит: «Он поднял на новый уровень понимание радиотехники... Он был человеком широчайшей эрудиции... Это был ученый в самом высоком смысле этого слова, который всего себя отдавал науке и старался делиться своими знаниями бескорыстно со всеми окружающими».

В 1949 году в Горьковском индустриальном институте специальность и кафедра «Радиотехника» были на электротехническом факультете, из преподавателей только один В.Я. Сморгонский имел степень кандидата технических наук. Дмитрий Васильевич возглавил кафедру и начал работу по подготовке научных кадров. Была организована аспирантура по радиотехническим специальностям, началось издание научных трудов,

регулярное проведение научных семинаров. Был резко увеличен объем научной работы, вырос ее уровень. Совету электротехнического факультета под председательством профессора Д.В. Агеева было предоставлено право присуждения степени кандидата технических наук. Первая защита диссертации, подготовленной самостоятельно Г.В. Глебовичем, состоялась в 1951 году. Свои новые идеи и научные результаты Дмитрий Васильевич щедро раздавал аспирантам-выпускникам института и соискателям-преподавателям. Каждому из них он отдал частичку своего таланта, он помогал довести идеи до завершения. Он отказывался от совместных с учениками публикаций, доверяя им самостоятельно публиковать полученные под его руководством результаты. Наибольшее внимание и поддержку он оказал первым из них: В.В. Маланову, Л.А. Моругину. В дальнейшем они стали докторами наук и заведующими кафедрами. В 1952 году по настоянию профессора Д.В. Агеева был восстановлен радиотехнический факультет. Он был учрежден еще в 1936 году, но в 1938 году преобразован в специальный факультет, а в 1947 – в электротехнический. Потребность в радиоинженерах была велика, на электротехническом факультете были две группы «электриков» и в 1950 году шесть, а в 1951 – восемь групп «радиостов». На возрожденном факультете из разросшейся кафедры радиотехники выделились две новые кафедры: радиоприемных устройств (зав. кафедрой профессор Д.В. Агеев) и радиопередающих устройств (зав. кафедрой доцент В.Я. Сморгонский). С 1955 года по 1959 профессор Д.В. Агеев был деканом факультета. Со времени восстановления радиотехнический факультет являлся самым большим в институте, сейчас это относится к ИРИТу в университете.

Педагогическая и учебно-методическая деятельность.

Гениальные люди гениальны во всем. Педагогическая и учебно-методическая деятельность профессора Д.В. Агеева, также, как и научная, была выдающейся. Главные ее результаты таковы.

Теория и метод лекционного обучения. Заведующий кафедрой, ведущий лектор двух радиотехнических дисциплин профессор Д.В. Агеев имел многолетний опыт преподавания в вузах страны. Он понимал, в чем состоят недостатки пришедшей из глубины веков лекционной системы обучения, которую был вынужден использовать сам, понимал необходимость ее кардинального улучшения на научной основе. Он не мог остаться равнодушным к этому и направил свои творческие способности, интеллект и эрудицию на совершенствование системы обучения. Этой творческой работой профессор Д.В. Агеев занимался более 10 лет. В развитой им теории обучения получены научно обоснованные ответы на многие важные для обучения вопросы. В частности, что такое эффективность обучения, каковы объективные критерии ее качественной и количественной оценки, что такое метод обучения, как влияет выбор важнейших целей обучения и другие. На основе теории профессор Д.В. Агеев разработал активный метод лекционного обучения студентов технических вузов [5]. В существующей системе лекция служит для понимания и в некоторой мере запоминания изучаемого материала с оценкой на экзамене. При использовании активного метода цели обучения таковы. Во-первых, научить, точнее, приучить студентов самостоятельно изучать материал по учебникам и книгам, а не получать его от лектора в «разжеванном» виде. Во-вторых, главное, научить творчески мыслить, то есть применять знания с использованием учебников и книг для решения творческих задач. Лекция служит для решения таких задач с помощью лектора с пояснением всей последовательности логических умозаключений, приводящих к решению. В конце каждой лекции проводится контрольная работа, в течение 15-20 минут все студенты потока решают одну творческую задачу. На следующей лекции рассматривается правильное решение контрольной задачи. Работы проверяются и оцениваются. В случае неявки студент получает нулевой балл. Окончательная оценка определяется как среднее между средним баллом за контрольные работы и баллом за

экзамен. На экзамене все студенты группы в течение полутора часов решают две одни и те же задачи, одну творческую, другую с преобладанием фактического материала. Для реализации такого оптимального варианта активного метода требуется увеличение времени на самостоятельную работу студентов и сокращение времени лекций. Пока этого перераспределения времени нет, используется несколько упрощенный квазиоптимальный вариант метода. Только половина фактического материала дисциплины изучается студентами самостоятельно, наиболее важная часть изучается на лекции. На лекции решается одна творческая задача. Контрольные работы и экзамен проводятся, как и при оптимальном варианте.

Активный метод лекционного обучения имеет следующие несомненные преимущества. Цели обучения соответствуют требованиям будущей работы студентов. В полной мере реализуются самостоятельность, активность и систематичность работы студентов, почти стопроцентным становится посещение лекций. Создается постоянная обратная связь между студентами и лектором, возможность воспитательного действия на студентов. Постоянно решая задачи, студенты фактически готовятся к экзамену, приобретают уверенность в своих силах, экзамен проходит в спокойной обстановке.

В течение четырех лет профессор Д.В. Агеев при чтении лекций по курсу «Усилительные устройства» в порядке эксперимента применял квазиоптимальный вариант активного метода обучения в учебном потоке из пяти групп. За семестр студенты выполняли 33 контрольные работы и решали две задачи на экзамене. Все преимущества нового метода были полностью подтверждены на практике.

Дмитрий Васильевич в центральных журналах опубликовал две статьи о разработанном им активном методе лекционного обучения [6, 7]. В 1984 году он прочел для преподавателей вузов города курс лекций объемом 40 часов на тему «Проблемное обучение и активный лекционный метод». Созданную им теорию и разработанный на ее основе активный метод обучения профессор Д.В. Агеев подробно изложил в подготовленной к изданию рукописи монографии [5]. В качестве эпиграфа к ней взято имеющее глубокий смысл утверждение «Примененье – мать ученья», созвучное общеизвестному. В рукописи приведены доказательства того, что эффективность разработанного активного метода является наиболее высокой по сравнению с другими принципиально возможными методами обучения. В ней дан детальный анализ результатов практического применения нового метода, приведены статистические данные об оценках контрольных работ и экзаменов 400 студентов. Приведены также результаты проведенной дискуссии, то есть самые острые вопросы и критические замечания, высказанные при обсуждении рукописи, и ответы автора на них. На подробный автореферат рукописи монографии были получены положительные отзывы ведущих кафедр вузов городов Москвы, Ленинграда, Минска, Харькова, Одессы, Горького. В отзывах отмечалось соответствие активного метода требованиям по повышению качества подготовки молодых специалистов и желательность скорейшего опубликования монографии. Дмитрий Васильевич предвидел и ощутил на себе трудности внедрения нового метода. Любая хорошая теория может быть дискредитирована плохой практикой ее осуществления. В данном случае препятствиями, которые необходимо преодолеть, являются психологический барьер – установившаяся традиция, отсутствие задачников с детальным рассмотрением методов решения творческих задач и недостаток времени для самостоятельной работы студентов. К сожалению, профессор Д.В. Агеев не смог опубликовать монографию, в 1988 году он был вынужден уйти на пенсию.

Приведенные выше результаты получены Д.В. Агеевым впервые в мире и за 25 лет до появления программы о переходе на компетентностно-ориентированное обучение. Эти результаты развития теории и практики высшего технического образования имеют мировой уровень и являются главными достижениями НГТУ в методической работе. Они представляют несомненный интерес для преподавателей технических вузов, монография Д.В. Агеева заслуживает опубликования в университете.

Создание радиотехнического факультета. Активизация научной работы и подготовка научно-педагогических кадров, начатые профессором Д.В. Агеевым сразу после приезда в город Горький, продолжались в течение всего времени его работы. В результате из 55 подготовленных под его научным руководством кандидатов технических наук 32 – преподаватели факультета. Из них в дальнейшем семь стали докторами наук и пять –заведующими кафедрами. Тем самым Дмитрий Васильевич внес самый существенный вклад в становление и развитие радиотехнического факультета, который трудно переоценить. В течение примерно 25 лет он сам и его ученики определяли уровень научной и учебно-методической работы, который непрерывно повышался. По значению новых научных идей и изобретений, по количеству и качеству научных работ, по числу монографий и учебных пособий, выпущенных центральными издательствами большими тиражами, по квалификации научно-педагогических кадров, по качеству подготовки радиоинженеров наш факультет стал в один ряд со старейшими радиотехническими факультетами страны. Фактически Дмитрий Васильевич создал факультет. Для сохранения памяти о гениальном ученом, внесшим фундаментальный вклад в отечественную и мировую науку и создавшим радиотехнический факультет, наследник факультета – ИРИТ – должен носить имя профессора Д.В. Агеева.

Кафедра «Радиоприемные устройства». Со времени создания кафедры 36 лет профессор Д.В. Агеев был заведующим кафедрой, а все преподаватели были его учениками. С начала семидесятых годов прошлого века в течение многих лет это была единственная в институте кафедра, все преподаватели которой имели ученые степени. В дальнейшем четыре преподавателя из восьми в указанной последовательности стали профессорами: В.Ф. Рябков, А.В. Зенькович, В.Н. Ногин, Н.Б. Догадин. Последний начал работу на кафедре в должности лаборанта. На кафедре активно проводилась научная работа, ее результаты имели признание в нашей стране и за рубежом. В первую очередь это фундаментальные результаты теоретических исследований профессора Д.А. Агеева. Некоторые из них были доведены до внедрения в ходе хозяйственных работ, проводимых на кафедре под его руководством. В центральных издательствах вышли в свет пять научных монографий сотрудников кафедры, было сделано 10 докладов на пленарных заседаниях Всесоюзных конференций. Было опубликовано более 260 научных статей, получено 107 авторских свидетельств на изобретения [8].

На международных зарубежных конференциях сотрудниками кафедры было сделано семь докладов, частично опубликованных в трудах конференций. Разработанные на кафедре приборы демонстрировались на международных выставках в Москве и Женеве. Начиная с 1977 года кафедра под руководством профессора Д.В. Агеева провела в институте четыре Всесоюзных конференции по новой технике радиоприемных устройств. Автор данной статьи с 1978 по 1988 год был организатором и научным руководителем тематического семинара «Измерения в технике радиоприема», проводимого два раза в год в Москве в павильоне «Физика» АН СССР. Уникальным является годовой экономический эффект внедрения результатов его докторской диссертации. В радиоизмерительной отрасли (15 типов серийно выпускаемых приборов) и в отрасли радиовещания (11 разработанных на кафедре установок) он в соответствии с актами внедрения составил 10,6 млн рублей, то есть в современных ценах около 1 млрд рублей. Разработанные автором статьи методы были включены в программу дисциплины «Электрорадиоизмерения» для вузов по специальности «Радиотехника».

Дмитрий Васильевич был ведущим лектором дисциплин «Радиоприемные устройства» и «Усилительные устройства», он написал два 100-страничных учебных пособия для спецкурсов по ним. Преподавателями кафедры по этим дисциплинам были написаны соответствующие программам учебные пособия, опубликованные центральными издательствами большими тиражами (1972 год – 30000 экз., 1992 – 3600 экз.). Тиражом 40000 экз. в 1985 году было издано учебное пособие по

электрорадиоизмерениям, в которое со ссылкой на автора вошли разработанные на кафедре методы измерения. В дальнейшем по этой дисциплине в институте был издан комплекс учебно-методических материалов объемом 23 печатных листа. В связи с наличием в городе головного приборостроительного института страны лаборатория радиоизмерений кафедры по глубине охвата тем дисциплины и числу используемых студентами приборов существенно превосходила аналогичные лаборатории всех вузов. По всем дисциплинам кафедры было издано 17 внутривузовских пособий объемом пять и более печатных листов. Дмитрием Васильевичем были опубликованы две монографии, 52 научные работы, получено 19 авторских свидетельств на изобретения.

Изложенные выше результаты научной и учебно-методической работы кафедры под руководством профессора Д.В. Агеева в значительной мере обеспечивали факультету в течение многих лет первое место среди факультетов института.

Дмитрий Васильевич был руководителем научного семинара факультета, редактором научных трудов, председателем диссертационного совета, членом аналогичного совета Горьковского госуниверситета, председателем областного совета НТО РЭС им. А.С. Попова, членом методической комиссии министерства, членом редколлегии журнала «Радиоэлектроника», членом совета издательства «Советское радио».

За многолетнюю плодотворную научную и педагогическую деятельность профессор Д. В. Агеев был награжден Орденом Трудового Красного Знамени, ему было присвоено звание «Заслуженный деятель науки и техники РСФСР», а также звание «Почетный радист СССР».

Выйдя в 1988 году на пенсию, Дмитрий Васильевич продолжал глубокие научные исследования. У него было много новых идей и большие планы, которым не суждено было сбыться. Он скончался 31 июля 1997 года на 87-ом году жизни.

Неиссякаемое стремление познать непознанное, неустанный труд, могучий интеллект, всепроникающая глубина анализа и неумолимая логика, глубокая научная и человеческая порядочность, доброжелательность и скромность — черты этого великого ученого и педагога.

Материалы о жизни и деятельности профессора Д.В. Агеева приведены в энциклопедической монографии профессора М.А. Быховского [9] и на сайте биографий великих людей всех времен и народов <http://great-people.ru/>. Более подробные сведения имеются в его мемориальном кабинете (аудитория 1323) на кафедре информационных радиосистем НГТУ.

Библиографический список

1. Zade L.A., Miller K.S. Fundamental Aspects of Multiplexing. Proc. IRE, 1952, №9.
2. Shannon C.E. A Mathematical Theory of Communication //BSTJ, 1948, № 3.
3. Wiener N. Cybernetics. -N.Y.: Wiley, 1948.
4. Агеев Д.В. Теория селекции и проблема пропускной способности «эфира» / Д.В. Агеев // НГТУ им. Р.Е. Алексеева. - Н. Новгород, 2017. – 231с.
5. Агеев Д.В. Активный метод лекционного обучения студентов технических вузов / Д.В. Агеев // -146 с. (Рукопись, архив кафедры)
6. Агеев Д.В. Активный лекционный метод / Д.В. Агеев // ЭКО - 1983. №1. С. 149-153.
7. Агеев Д.В. Внимание – активный метод / Д.В. Агеев // Студенческий меридиан. 1984. № 3.
8. Ногин В.Н. Ранняя история радиофакультета и история кафедры «Радиоприемные устройства» Горьковского политехнического института (ныне НГТУ) / Ногин В.Н. // История разработок и производства радиоэлектронных средств с 1915 года на предприятиях Нижнего Новгорода. Под ред. Е.Л. Белоусова, Б.Д. Увяткина, В.Л. Ягодкина - Н. Новгород. 2001. С. 397-421.
9. Быховский М.А. Пионеры информационного века / М.А. Быховский // История развития теории связи. Сер. История электросвязи и радиотехники. М.: ЗАО «РИЦ» «Техносфера», 2006. – 376 с.

A.V. Zenkovich

**ON THE 110TH ANNIVERSARY OF GREAT SCIENTIST AND TEACHER
PROFESSOR DMITRY VASILYEVICH AGEYEV**

Nizhny Novgorod State Technical University n.a. R.E. Alekseev

Professor D.V. Ageyev, the founder of the theory of communication and transmission of information, created more than 10 new theoretical directions of research in radio engineering. Results of his theoretical investigations were pioneering and fundamental. Professor D.V. Ageyev trained 55 candidates of science (Ph.D.), 14 of them became then Doctors of Science. He created Gorky scientific school in radio engineering. In fact, he created the radio engineering department and made it comparable with the best departments of the country. His chair took the leading position at the institute. Professor D.V. Ageyev also created a theory of education, he developed the active method of delivering lectures and applied it practically for 4 years.

Keywords: founder of communication theory, scientific school, theory and active method of education, creating the department.

ОБЗОР ТРЕНДОВ В ТЕХНОЛОГИЯХ ВЫСОКОПРОИЗВОДИТЕЛЬНЫХ ВЫЧИСЛЕНИЙ, ИСКУССТВЕННОГО ИНТЕЛЛЕКТА И ЦЕНТРОВ ОБРАБОТКИ ДАННЫХ В КОНТЕКСТЕ РАЗВИТИЯ ЦИФРОВЫХ ЭКОСИСТЕМ

¹*Волжский государственный университет водного транспорта,*
²*Исследовательский центр Huawei, Нижний Новгород*

В контексте развития цифровых экосистем рассматриваются тренды внедрения технологий высокопроизводительных вычислений, искусственного интеллекта и центров обработки данных.

Ключевые слова: центры обработки данных, искусственный интеллект, тренды внедрения цифровых технологий

В течение последних десяти лет на мировом ландшафте базовых цифровых технологий наблюдаются принципиальные изменения. От доминирования Intel с архитектурой x86 и NVidia с решением CUDA происходит существенное смещение в сторону альтернатив в виде конкурентоспособных ARM решений компании ARM Limited.

Существенное количество инноваций обнаруживает за этот период и индустрия центров обработки данных (ЦОД). От помещений с компьютерным, информационным, и телекоммуникационным (ИКТ) оборудованием происходит переход к облачным ЦОДам.

Быстрое развитие новых технологий таких, как Искусственный интеллект (ИИ) [1, 2], Большие данные (БД) и 5G способны оказать влияние на большинство направлений науки и техники, социума и хозяйственной деятельности. Соответственно спрос на ЦОДы значительно возрастет. Ниже в статье рассматриваются основные тренды во внедрении вышеупомянутых технологий.

1. С технической точки зрения Intel стремится к максимальной производительности ядра при относительно небольшом их количестве в процессоре. NVidia «смотрит» в сторону большого количества специализированных вычислительных ядер в процессоре. Решения ARM занимают промежуточное положение, с относительно большим по сравнению с решениями Intel количеством ядер и большей производительностью на поток. Обращают на себя внимание следующие аспекты, которые сказываются на конкурентоспособности технологии ARM:

а) В актуальном рейтинге суперкомпьютеров появился суперкомпьютер Fugaku на базе ARM. Причём, кластер занял первое место по производительности в списке и показывает 0.5 эксафлопа, обгоняя второй в списке более чем в два раза [3].

б) Компания Apple выпустила версию Macbook pro на ARM.

в) Технология ARM открыта для использования.

2. В области ИИ выделяются следующие области, расположенные по убыванию применимости и возрастанию консервативности [1, 4]. Это:

- розничная торговля;
- кибербезопасность;
- образование;
- транспортные средства;
- медицина;
- технические сегменты (тренировка большего количества гиперпараметров, перенос вычислений ближе к источнику данных и дальше от сервера, канал поставки данных становится большей проблемой, чем скорость вычислений);
- академические;

- геополитические (по широте использования выделяется КНР; в РФ разрабатываются законы, регламентирующие применение ИИ; с учетом позиций защиты прав человека в странах ЕС наблюдается сдерживание применения технологий ИИ).

3. Рассмотрим тренды в области ЦОД [1].

3.1. Высокая плотность. Производительность центрального процессора и серверов постоянно растёт по мере развития технологий. Ввиду роста спроса на приложения для ИИ требуемая вычислительная мощность постоянно возрастает. Чтобы сбалансировать эффективность и затраты ЦОДы должны развиваться в направлении высокой плотности. В настоящее время средняя потребляемая мощность в ЦОДе составляет от 6 до 8 кВт на стойку. Ожидается, что к 2025 году этот показатель будет располагаться в интервале от 15 до 20 кВт.

3.2. Масштабируемая архитектура. Как правило, жизненный цикл ИТ-устройств колеблется от 3 до 5 лет, а удельная производительность удваивается каждые 5 лет. Однако, жизненный цикл инфраструктуры ЦОД лежит в интервале от 10 до 15 лет. Таким образом, ЦОДы должны поддерживать эволюцию ИТ-устройств в течение 2 - 3 поколений. Это потребует масштабируемое расширение и поэтапные инвестиции для оптимизации капитальных затрат на протяжении всего жизненного цикла ЦОД. Кроме того, из-за разнообразия ИТ-услуг, ЦОДы должны поддерживать гибридное развертывание устройств с различным энергопотреблением.

3.3. Минимизация воздействия на окружающую среду. В настоящее время энергопотребление ЦОДов составляет 3% от мирового потребления энергии. По существующим оценкам к 2025 году общая потребляемая мощность ЦОДов достигнет более 1000 ТВтч. Важнейшие вызовы для построения ЦОДов будущего это энергосбережение, сокращение выбросов и снижение эксплуатационных расходов.

Снижение коэффициента эффективности использования энергии (PUE) и строительство «зеленых» ЦОДов неизбежны. В мире распространяется тенденция использования чистой энергии и отходящего тепла, а современная ИТ-индустрия стремится экономить ресурсы на протяжении всего жизненного цикла ЦОДа. Предполагается, что среднее значение PUE новых ЦОДов в Китае снизится до 1,1 в ближайшие пять лет.

3.4. Быстрое развертывание. В сфере Интернета количество запросов постоянно возрастает. Соответственно требования к данным и трафику на стороне представления услуг также увеличиваются. Возникает необходимость оперативного развертывания новых ЦОДов. С другой стороны, ЦОД из системы поддержки превращается в производственную систему. Более быстрое развертывание означает более быструю выгоду. Средний срок запуска типового ЦОДа сегодня колеблется от 9 до 12 месяцев, а в будущем он будет сокращен до величины менее 6 месяцев.

3.5. Полная цифровизация и применение ИИ. ЦОДы неизбежно будут эволюционировать в сторону применения интеллектуальных технологий. По мере постоянного совершенствования технологий Итернета Вещей и ИИ ЦОДы постепенно перейдут от цифровизации только в части эксплуатации и техобслуживания, энергосбережения и эксплуатации, к цифровизации и автоматическому управлению на протяжении всего жизненного цикла с точки зрения планирования, построения, эксплуатации и техобслуживания, а также оптимизации этих процессов.

3.6. Полностью модульная структура. Для решения проблем медленного строительства и высоких первоначальных инвестиционных затрат будет построено больше полностью модульных ЦОДов. Модульная структура будет эволюционировать от модульности на уровне компонентов к модульной архитектуре и модуляризации помещения оборудования, что в конечном итоге приведет к полной модуляризации ЦОДа.

Полная модульная конструкция обеспечивает быстрое развертывание, простую эксплуатацию и техобслуживание, а также высокую энергоэффективность.

3.7. Упрощенная архитектура электропитания. Система распределения питания традиционного ЦОДа является сложной и фрагментированной, занимает большую площадь и в ней сложно найти неисправности, проводить ремонт и техническое обслуживание. Упрощенная архитектура источника питания сокращает время преобразования энергии, уменьшает расстояние и занимаемую площадь источника питания, повышает коэффициент использования пространства и энергоэффективность системы в целом.

По сравнению со свинцово-кислотными батареями, литиевые батареи имеют преимущества с точки зрения занимаемой площади и срока службы. По мере снижения стоимости литиевых батарей, они все чаще будут широко использоваться в ЦОДах.

3.8. Объединение жидкостного и воздушного охлаждения. Увеличение применения непрямого испарительного охлаждения и уменьшение охлаждения водой. Графические и нейронные процессоры генерируют больше сценариев с высокой плотностью и соответствующим энергопотреблением, а также выделением тепла. Соответственно системы жидкостного охлаждения будут всё более и более популярными. Тем не менее, некоторые услуги хранения и вычисления все еще находятся в сценариях с низкой плотностью.

Для быстрой адаптации к неопределенным требованиям ИТ-услуг в будущем решения охлаждения процессоров должно быть совместимо с системами воздушного и жидкостного охлаждения. Кроме того, сложная архитектура системы охлаждения воды затрудняет быстрое развертывание и упрощается установкой системы охлаждения с помощью непрямого испарения с модульной архитектурой. Благодаря полному использованию природных ресурсов охлаждения, энергопотребление системы охлаждения будет значительно сокращено. В районах с подходящим климатом система охлаждения водой будет постепенно заменена системой непрямого испарительного охлаждения.

3.9. Динамическая связь между битами и ваттами. Снижение PUE не означает, что общее энергопотребление ЦОД оптимально. Вместо того чтобы сосредоточиться только на электрическом оборудовании, необходимо оценить и оптимизировать энергопотребление ЦОДа в целом.

Благодаря инновациям всего технологического стека, который включает оборудование, наборы микросхем, данные и облачные технологии, биты и ватты будут работать совместно для достижения динамической экономии энергии и оптимальной энергоэффективности всей системы.

3.10. Надежность. По мере того, как средства ЦОДов становятся более интеллектуальными, угрозы сетевой безопасности будут только возрастать.

ЦОД должен обладать следующими пятью функциями. Это устойчивость, безопасность, конфиденциальность, надежность и доступность для предотвращения атак и угроз от окружения и недобросовестного персонала, включая угрозы сетевого вторжения.

Библиографический список

1. Yiwei Wang, Lei Huang, Siwen Jiang, Yifei Wang, [Jun Zou](#), Hongguang Fu, and [Shengyong Yang](#) Capsule Networks Showed Excellent Performance in the Classification of hERG Blockers/Nonblockers. <https://www.frontiersin.org/articles/10.3389/fphar.2019.01631/full>
2. Jifeng Dai, Haozhi Qi, Yuwen Xiong, Yi Li, Guodong Zhang, Han Hu Yichen Wei Deformable Convolutional Networks, 2017. <https://arxiv.org/pdf/1703.06211.pdf>, <https://top500.org/lists/top500/2020/11/>, <https://www.huawei.com/en/news/2020/2/huawei-top10-trends-datacenter-facility-2025>

S.V. Pavlov^{1,2}, Y.S. Fedosenko¹

TRENDS OVERVIEW IN HIGH PERFORMANCE COMPUTING TECHNOLOGIES, ARTIFICIAL INTELLIGENCE AND DATA PROCESSING CENTERS IN THE CONTEXT OF DIGITAL ECOSYSTEMS DEVELOPMENT

¹*Volga State University of Water Transport, Nizhny Novgorod*

²*Huawei Research Center, Nizhny Novgorod, Russia*

In the context of the development of digital ecosystems, trends in the introduction of high-performance computing technologies, artificial intelligence and data centers are considered.

Keywords: data centers, artificial intelligence, trends in the introduction of digital technologies.

УДК 004.05

К.А. Батенков

ОБОБЩЕННЫЙ ВЗГЛЯД НА КАЧЕСТВО ОБСЛУЖИВАНИЯ В ТЕЛЕКОММУНИКАЦИЯХ

Академия ФСО России, г. Орёл

В работе представлен обобщенный взгляд на качество обслуживания в телекоммуникациях, предполагающий его рассмотрение с четырех точек зрения. Показано, что в общем случае модель рассматривает два измерения: взаимоотношения между пользователем услуги и поставщиком услуги; ожидания и достигнутый уровень качества обслуживания для обеих сторон – пользователя и поставщика услуг.

Ключевые слова: сети связи, качество обслуживания, модель четырех точек зрения, приложение.

Введение

Существуют принципиальные различия между качеством функционирования и обслуживания [1]. Практически всегда возможно выделить набор параметров, относящихся к различным аспектам услуг и выбрать соответствие между их значениями и интервалом наблюдения. Это позволяет оценить только технические характеристики в соответствии с используемыми критериями, но не характеристику услуги в целом.

Следует подчеркнуть, что ключевой особенностью любой телекоммуникационной услуги является необходимость использования как минимум двух окончаний на сторонах пользователей. В результате обе стороны могут оказывать влияние на качество услуги, а, следовательно, их учет при анализе чрезвычайно важен. Кроме того, поскольку телекоммуникационные услуги являются высокотехнологичными и воспринимаются пользователями комплексно, то психологические аспекты оказываются критичными при оценке качества услуг. В итоге удовлетворенность потребителей – это результат восприятия качества, вытекающий из сравнения их ожиданий с реально наблюдаемыми фактами.

1. Требования к качеству телекоммуникационной услуги

При организации телекоммуникационных услуг используются разнообразные технологии и средства, показатели и метрики (KQI – key quality indicator) для которых должны соответствовать требованиям к уровню обслуживания (SLO – Service Level Objectives). Для констатации факта достижения требуемого уровня обслуживания провайдеры применяют средства для анализа ключевых показателей эффективности (KPI – key performance indicator).

Требования к уровню обслуживания (SLO) – специализированные требования пользователей, выражаемые в форме поддающихся оцениванию характеристики уровня обслуживания.

Ключевой показатель качества (KQI) – параметр, включающий определенные составляющие производительности услуги и продукта, значимые для пользователей. Ключевой показатель качества обычно объединяет несколько ключевых показателей эффективности, промежуточно вычисляемых компонент обычно на основе KPI, другие KQI от одного или даже нескольких операторов и непосредственные измерения, использующие математические соотношения, называемые функциями оценки KQI.

Ключевой показатель эффективности (KPI) – параметр, включающий определенные составляющие производительности средств, включая поставщика и услуги,

измеряемый либо напрямую, либо в форме иерархии. Ключевой показатель эффективности значим для оператора, но не обязателен для пользователя.

Требования к уровню обслуживания в общем случае определяются в терминах возможных значений, или уровня обслуживания, целей, интервала, средств и времени измерений. Например, 90 % звонков в службу поддержки, обслуженные менее чем за 20 секунд, измеренные на интервале одного месяца, зарегистрированные системой автоматического распределения вызовов. Соглашение между пользователем и провайдером относительно требований к уровню обслуживания формулируется в договоре об уровне обслуживания (SLA – service level agreement). Регламентированы основные свойства к требованиям к уровню обслуживания [1]: достижимость, повторяемость, оцениваемые, понятность, значимость, контролируемость, экономическая целесообразность, непротиворечивость. Показатели, относящиеся к требованиям к уровню обслуживания, измеряются и контролируются для подтверждения соответствия указанного в договоре уровня обслуживания.

Когда сервис является частью универсальной (общедоступной) услуги, регулятор выступает гарантом соответствия телекоммуникационной услуги правовым нормам. В случае сервиса, описываемого в договоре об уровне обслуживания между пользователем и провайдером, параметры фиксируются в данном договоре. В обоих вариантах для устранения недопонимания методы измерения и контроля должны быть определены и доступны для всех заинтересованных сторон.

Понятие качества применительно к телекоммуникациям является достаточно сложным и противоречивым термином вследствие отсутствия в настоящее время четкого набора параметров, которые бы характеризовали тот или иной аспект качества. Существуют лишь отдельные рекомендации ИТУ и технические описания ETSI, регламентирующие наборы параметров для определенных видов услуг или же точек сети.

Тем не менее, существует общепринятые и широко используемые модели качества в телекоммуникациях, определяющие качество как для различных точек измерения в телекоммуникационной системе, так и для различных этапов обслуживания пользователей.

2. Модель качества обслуживания четырех точек зрения

Качество обслуживания (QoS – quality of service) – совокупность характеристик услуги электросвязи, которые имеют отношение к ее возможности удовлетворять установленные и предполагаемые потребности пользователя услуги [2].

В рекомендации ИТУ-T G.1000 [3] определена модель качества обслуживания четырех точек зрения. В общем случае модель рассматривает два измерения [4]:

- взаимоотношения между пользователем услуги и поставщиком услуги;
- ожидания и достигнутый уровень качества обслуживания для обеих сторон – пользователя и поставщика услуг.

Ожидаемое и воспринимаемое качество обслуживания с точки зрения пользователя представляется в достаточно наглядных и понятных терминах, в то время поставщик использует техническую терминологию для описания предлагаемого и достигнутого уровня качества обслуживания [3]. Таким образом, качество необходимо рассматривать с четырех точек зрения (рисунок 1):

- требования клиента к качеству обслуживания;
- предложения (планируемое, целевое) качества обслуживания поставщика услуг;
- предоставленное или достигнутое качество обслуживания;
- клиентский рейтинг качества обслуживания.

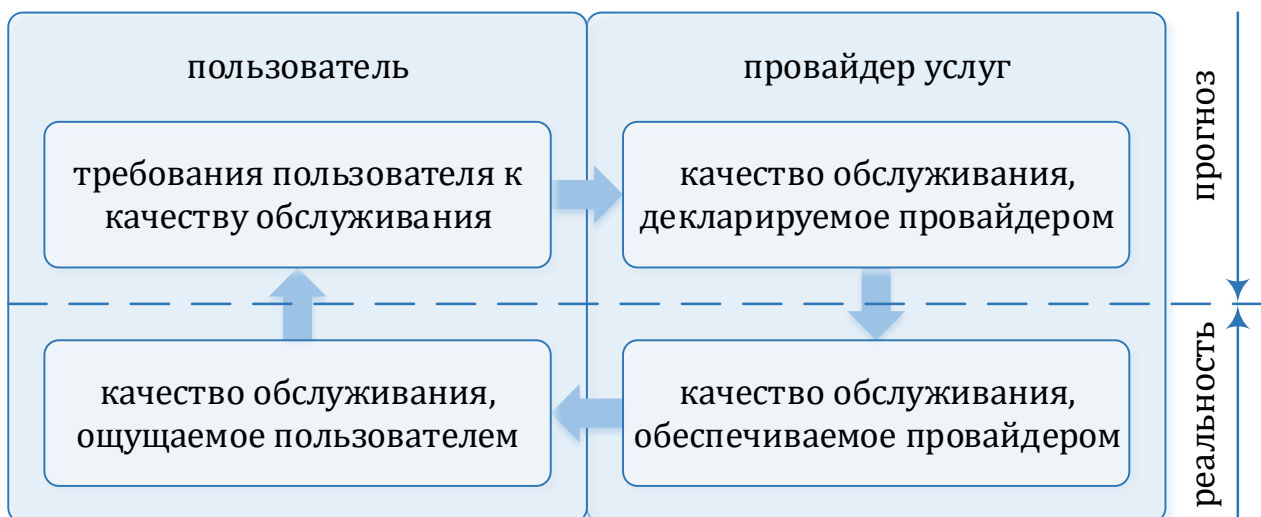


Рис. 1. Модель качества обслуживания четырех точек зрения

3. Требования пользователя к качеству обслуживания

Требования пользователя к качеству обслуживания определяют уровень качества, необходимый для конкретной услуги, выраженный нетехническим языком. Клиенту безразличны способ предоставления конкретной услуги и структура сети. Он заинтересован только в достижении сквозного качества обслуживания. С точки зрения заказчика, качество обслуживания выражается параметрами со следующими свойствами:

- отражают ощущаемые пользователем эффекты, а не причины их обуславливающие;
- не зависят, по их определению, от структуры сети;
- учитывают все нюансы услуги с точки зрения клиента;
- гарантируются поставщиком услуг, иногда в форме договора;
- описываются в независимых от сети терминах и создают общий язык, понятный как пользователю, так и поставщику услуг.

Особый интерес представляет рекомендация ITU-T Rec. G.1010 [5], описывающая требования к качеству обслуживания конечных пользователей для типовых приложений, которые используются поставщиком услуг для планирования уровня обслуживания.

4. Качество обслуживания, декларируемое провайдером

Качество обслуживания, декларируемое провайдером – это заявление об уровне качества, который, как ожидается, будет предоставлен клиенту поставщиком услуг. Уровень качества выражается значениями параметров качества обслуживания. Основное предназначение подобного вида качества обслуживания – формирование исходных данных при планировании и для соглашения об уровне обслуживания (SLA). Каждая услуга имеет свой собственный набор параметров качества обслуживания, например, соответствующих классам сетей IP (internet protocol – протокол интернета) рекомендации ITU-T Y.1540 [6]. Поставщик услуг может выражать предлагаемое качество обслуживания в нетехнических терминах в интересах клиентов и в технических терминах для использования в коммерческих целях.

Например, предложение поставщика услуг с пользовательской точки зрения может формулироваться как готовность базовой телефонной услуги должна составлять 99,95% за год с не более чем пятнадцатиминутным перерывом в любой момент времени и не более чем тремя перерыва в течение года.

5. Качество обслуживания, обеспечиваемое поставщиком услуг

Качество обслуживания, обеспечиваемое поставщиком услуг, – это заявление об уровне качества, фактически достигнутом и предоставленном клиенту [7–9]. Оно выражается значениями параметров, идентичными параметрам предлагаемого качества обслуживания. Их сравнение позволяет сделать вывод о фактически достигнутом уровне обслуживания. Эти показатели производительности эффективности обобщаются за определенные периоды времени, например, за предыдущий месяц.

Например, заключение поставщика услуг может быть следующим: достигнутая готовность за предыдущий квартал составила 99,95% с пятью перерывами обслуживания, из которых один длился 65 минут. Достигнутое или предоставленное качество обслуживания анализируется производителями, иногда регуляторами в интересах клиентов.

6. Качество обслуживания, воспринимаемое пользователями

Качество обслуживания, воспринимаемое пользователями, – это утверждение, отражающее уровень качества, который с их точки зрения они получили. Воспринимаемое качество обслуживания выражается, как правило, в терминах степени удовлетворенности, а не в техническом смысле. Воспринимаемое качество обслуживания оценивается на основе опросов клиентов и их комментариев об уровне обслуживания. Воспринимаемое качество обслуживания используется поставщиком услуг для определения удовлетворенности клиентов качеством услуг. Например, клиент может констатировать, что в неприемлемом количестве случаев возникли трудности с установлением соединения, и может оценить качество на два балла по пятибалльной шкале. В идеальном случае должно быть полное соответствие между предоставленным воспринимаемым качеством обслуживания.

В качестве логической отправной точки можно рассматривать требования пользователя к качеству обслуживания, например, приведенные в рекомендации ITU-T G.1010 [5]. Набор требований клиента к качеству обслуживания может рассматриваться изолированно, поскольку предполагается их достижение. Подобные требования – исходные данные для определения поставщиком услуг предлагаемого или планируемого качества обслуживания [10, 11]. Поставщик услуг не всегда может быть в состоянии предложить клиентам запрашиваемый уровень качества обслуживания. Такие факторы как стоимость качества, стратегические бизнес-цели поставщиков услуг, аттестация ("лучший в своем классе") и другие влияют на уровень предлагаемого качества. Требования заказчика также влияют и на системы мониторинга, разворачиваемые с целью проверки достигаемого качества обслуживания. Сочетание взаимосвязанных точек зрения качество обслуживания образует основу практического и эффективного управления качеством услуг. Одним из признаков развития провайдера обычно считают схожесть всех четырех точек зрения для определенной услуги.

Библиографический список

1. EG 202 009-1. User Group; Quality of telecom services; Part 1: Methodology for identification of indicators relevant to the Users. – v. 1.3.1. – 2014–12. – Sophia Antipolis : ETSI, 2014. – 49 p.
2. Rec. E.800. Definitions of terms related to quality of service. – 2008–09. – Geneva : ITU-T, 2009. 30 p.
3. Rec. G.1000. Communications quality of service: a framework and definitions. – 2001–11. – Geneva : ITU-T, 2011. – 16 p.
4. Rec. E.804. QoS aspects for popular services in mobile networks. – 2014–02. – Geneva : ITU-T, 2014. – 446 p.
5. Rec. G.1010. End-user multimedia QoS categories. – 2001–11. – Geneva : ITU-T, 2002. – 18 p.

6. Rec. Y.1540. Internet protocol data communication service – IP packet transfer and availability performance parameters. – 2011–03. – Geneva : ITU-T, 2011. – 52 p.
7. Батенков К. А. Числовые характеристики структур сетей связи // Труды СПИИРАН. 2017. № 4 (53). С. 5–28.
8. Батенков К. А. Точные и граничные оценки вероятностей связности сетей связи на основе метода полного перебора типовых состояний // Труды СПИИРАН. 2019. Т. 18. № 5. С. 1093-1118.
9. Батенков К.А. Анализ надежности многополюсных сетей связи на основе метода минимальных сечений // В сборнике: Актуальные проблемы инфотелекоммуникаций в науке и образовании (АПИНО 2019) Сборник научных статей VIII Международной научно-технической и научно-методической конференции. В 4-х томах. Под. редакцией С.В. Бачевского. 2019. С. 108–112.
10. Батенков К. А. Анализ и синтез структур сетей связи по детерминированным показателям устойчивости / К. А. Батенков, А. А. Батенков // Труды СПИИРАН. 2018. № 3 (58). С. 128–159.
11. Батенков А. А. Методы формирования множеств состояний телекоммуникационных сетей для различных мер связности / А. А. Батенков, К. А. Батенков, А. Б. Фокин // Труды СПИИРАН. 2020. Т. 19. № 3. С. 644-673.

К. А. Batenkov

GENERALIZED VIEW OF QUALITY OF SERVICE IN TELECOMMUNICATIONS

Academy of Federal Guard Service of Russian Federation, Oryol

The paper presents a generalized view of the quality of service in telecommunications, suggesting its consideration from four points of view. It is shown that in the general case, the model considers two dimensions: the relationship between the service user and the service provider; expectations and the achieved level of service quality for both parties – the user and the service provider.

Keywords: communication, networks, quality of service, model of four points of view, application.

МЕТОДИЧЕСКИЙ ПОДХОД К ОБОСНОВАНИЮ ПРИОРИТЕТНЫХ НАПРАВЛЕНИЙ СОВЕРШЕНСТВОВАНИЯ РАДИОКОНТРОЛЯ С УЧЕТОМ ОЦЕНКИ ЕГО ДОСТОВЕРНОСТИ

*Военный учебно-научный центр Военно-воздушных сил
«Военно-воздушная академия имени профессора Н.Е. Жуковского и Ю.А. Гагарина»,
г. Воронеж*

В статье рассмотрены вопросы совершенствования выполнения мероприятий радиоконтроля с учетом оценки его достоверности. Оценка приоритетности направлений совершенствования радиоконтроля рассмотрена с позиции решения многокритериальной задачи.

Ключевые слова: Радиоконтроль, достоверность, коэффициент значимости параметров, эффективность радиоконтроля.

Одной из основных задач при оценке приоритетности направлений совершенствования радиоконтроля (РК) (далее – РК) объектов защиты Военно-воздушных сил (далее – ВВС) является выбор показателей и критериев оценки их приоритетности. Способом решения задачи при наличии нескольких критериев (так называемый критерий-вектор) является введение составного критерия вида:

$$W = a_1W_1 + a_2W_2 + \dots + a_nW_n, \quad (1)$$

где $W_1, W_2 \dots W_n$ – частные критерии, а $a_1, a_2, \dots a_n$ – весовые коэффициенты (или коэффициенты важности), сумма которых равна 1.

Поскольку каждый из частных критериев отражает разные и нередко противоречивые показатели (например, затраты ресурсов, которые желательно уменьшить, и боеготовность, которую необходимо повысить), имеющие свою размерность (руб. и т.д.), обобщенный критерий W может не иметь ясно выраженного физического смысла. Его следует рассматривать лишь как меру предпочтительности. Веса устанавливаются, как правило, экспертным путем [1].

С учетом того, что РК является многофакторным процессом, в качестве показателей целесообразно принять как целевые, так и функциональные показатели [1].

В качестве внешнего и, соответственно, приоритетного показателя эффективности РК целесообразно определить показатель предотвращенного ущерба за счет утечки информации, представляющей интерес для радиоразведки, или информативных материалов, по всем j -м каналам радиосвязи (далее – КРС) $W_{np.yu}$:

В качестве внешнего и, соответственно, приоритетного показателя эффективности РК целесообразно определить показатель:

$$W_{np.yu} = \sum_{j=1}^N W_{yu_j}, \quad (2)$$

где W_{yu_j} – предотвращенный ущерб за счет утечки информации, представляющей интерес для радиоразведки, или информативных материалов, по всем j -м каналам радиосвязи (далее – КРС).

Необходимость введения такого показателя определяется тем, что РК является специфической формой защиты информации от ее утечки по открытым каналам радиосвязи, используемых должностными лицами ВВС. При этом критерием для данного показателя целесообразно определить максимизацию значения предотвращенного ущерба $\max \{W_{np.ущ}\}$ за счет утечки по всем j -м КРС. По сути это и есть целевая функция РК, а $W_{np.ущ}$ – целевой показатель РК.

В результате принятия мер по исключению или снижению вероятности вскрытия смысловой информации, представляющей интерес для радиоразведки, по j -му КРС, происходит предотвращение потенциального ущерба $W_{np.ущj}$, который может быть нанесен в результате утечки информации за счет принятия контрмер по снижению эффективности образцов вооружения, военной и специальной техники, нормального функционирования военных объектов и других контрмер, предпринимаемых противником.

Значение $W_{np.ущj}$ можно определить, как:

$$W_{np.ущj} = W_{np.ущ максj} I_{jcp}, \quad (3)$$

где $W_{np.ущ максj}$ – максимальный предотвращенный ущерб, который может быть нанесен за счет утечки информации по j -му каналу; I_{jcp} – средневзвешенный (нормированный) коэффициент информативности j -го КРС, принимающий значения от 0 до 1 и характеризующий его информативность, определяемый следующим образом:

$$I_{jcp} = \frac{I_j}{I_{макс}}, \quad (4)$$

где I_j – коэффициент информативности j -го КРС; $I_{макс}$ – максимальный коэффициент информативности КРС.

Функциональные (внутренние) показатели эффективности РК определялись с учетом показателей, традиционно используемых для оценки качества технического контроля [2]:

- оперативность контроля;
- полнота контроля;
- достоверность контроля.

Применительно к РК показателем оперативности целесообразно определить показатель $O_{рк}$, который определяет время выполнения всех операций (мероприятий) РК, начиная от его планирования и заканчивая принятием мер по результатам выявленных в ходе проведения РК нарушений установленного порядка ведения обмена информацией (радиопереговоров):

$$O_{рк} = \frac{1}{T_{подг} + T_{пров}}, \quad (5)$$

где $T_{подг}$ – время подготовительного этапа РК;

$T_{пров}$ – время этапа проведения РК и обработки информации.

В качестве показателя полноты РК целесообразно определить показатель $\Pi_{рк} = F \{ \Pi_{крс}, \Pi_{ткрс}, \Pi_{инф} \}$, с учетом специфики РК, характеризующий:

полноту охвата РК информативных каналов радиосвязи (КРС) объектов защиты, $\Pi_{крс}$;

полноту времени РК каждого информативного КРС, $\Pi_{ткрс}$;

полноту обработки информации, полученной в ходе РК, $\Pi_{инф}$.

Кроме этого, в качестве функционального (внутреннего) показателя эффективности РК целесообразно использовать такой показатель, как достоверность РК, D_{pk} . Данный показатель ранее не учитывался в методическом подходе к обоснованию приоритетных направлений совершенствования РК.

В основу метода оценки достоверности РК может быть положен, с некоторыми допущениями, метод определения коэффициентов значимости параметров, используемый при контроле сложных систем и основанный на теории чувствительности [1,2]. Этот метод применяется для анализа систем со случайными параметрами в предположении малых дисперсий этих параметров по сравнению с их математическими ожиданиями. Гипотеза о малых дисперсиях параметров позволяет построить простые аппроксимирующие зависимости показателя эффективности РК от соответствующих параметров.

Использование данного подхода к рассматриваемой предметной области обусловлено тем, что современные измерительные системы обеспечивают точность измерений, позволяющих делать вывод о незначительности дисперсий параметров по сравнению со значениями самих параметров.

Предлагается следующее:

В качестве параметров методики расчета показателя достоверности РК на верхнем уровне рассматриваются радиотехнические параметры – мощность сигнала, частота сигнала и другие;

Рассматриваемые параметры являются независимыми (что «физически» не вызывает сомнений);

Связь параметров измерений с соответствующими параметрами нижестоящего уровня является аналитической.

Показатель достоверности РК D_{pk} должен учитывать [1,2]:

Разброс значений исследуемого параметра;

точность измерения (значимости определения) исследуемого параметра;

Влияние точности измерения (определения) единичного параметра на ошибку определения значения параметра РК.

Соответственно, оценку показателя достоверности РК D_{pk} целесообразно проводить в соответствии со следующим выражением:

$$D_{pk} = \sum_{k=1}^K R_k, \quad (6)$$

где R_k - коэффициент значимости k -го параметра. Чем больше коэффициент значимости, тем больший вклад в повышение достоверности дает технический контроль (измерение) данного параметра.

В соответствии с [2] коэффициент значимости k -го параметра R_k определяется следующим образом:

$$R_k = 1 - \sqrt{1 - \frac{B_k^2(D_k - D_k^*) + 2C_k^2(D_k - D_k^*)^2}{D_w}}, \quad (7)$$

где D_k и D_k^* - соответственно, априорное и апостериорное значения дисперсии k -го параметра; B_k и C_k - соответственно, центральная и вторая разностная производные исследуемой зависимости; D_w - дисперсия оценки значений исследуемой функции W . То есть, чем точнее измерение параметров, тем больше значение коэффициента значимости k -го параметра и тем выше достоверность РК в целом.

Расчет B_k и C_k производится по следующим формулам:

$$B_k = \frac{W(m_k + h_k) - W(m_k - h_k)}{2h_k}, \quad (8)$$

$$C_k = \frac{W(m_k + h_k) + W(m_k - h_k) - 2W_0}{h_k^2}, \quad (9)$$

где m_k - математическое ожидание k -го параметра;

W_0 - значение исследуемой функции при номинальных значениях параметров;

$$h_k = k_k + \delta_k, \quad (10)$$

где k_k - энтропийный коэффициент, представляющий собой отношение энтропийной погрешности к среднеквадратическому отклонению параметра (для равномерного закона распределения $k_k = 1,73$; для нормального $k_k = 2,07$);

δ_k - среднеквадратическое отклонение k -го параметра.

При оценке значимости параметров РК иногда важно знать вклад в достоверность РК не только отдельных параметров, но и их совокупности. Учитывая, что рассматриваемые параметры являются независимыми, формула для расчета значимости совокупности параметров принимает вид:

$$R_L = 1 - \sqrt{1 - \frac{\sum_{l=1}^L (B_l^2 (D_l - D_l^*) + 2C_l^2 (D_l - D_l^*)^2)}{Dw}}, \quad (11)$$

где L - общее количество рассматриваемых параметров.

С учетом приведенных показателей эффективности РК, а также их критериев, соотношение для расчета общего показателя эффективности РК, который будет использован для оценки приоритетности направлений совершенствования обеспечения РК в ВВС:

$$W_{pk} = a_1 W_{np.yu} + a_2 O_{pk} + a_3 \Pi_{pk} + a_4 D_{pk}. \quad (12)$$

Таким образом, предложен методический подход к обоснованию приоритетных направлений совершенствования РК, в котором, в отличие от ранее существующих методических подходов, учтен показатель достоверности РК, расчет которого основан на методе определения коэффициентов значимости параметров, используемом при контроле сложных систем и базирующемся на теории чувствительности.

Библиографический список

1. Жуков Г.П., Викулов С.Ф. Военно-экономический анализ и исследование операций. М. Воениздат, 1987. 80 с.
2. Технический контроль в машиностроении: Справочник проектировщика/ Под общ. редакцией В.Н. Чупырина, А.Д. Никифорова. М. Машиностроение, 1987. 512 с.

V.A. Bezryadin, R. A. Khvorov, B.T. Ryakhov

THE METHODOLOGICAL APPROACH TO JUSTIFICATION OF PRIORITY DIRECTIONS OF IMPROVING RADIO CONTROL, TAKING INTO ACCOUNT THE ASSESSMENT OF ITS RELIABILITY

Military Educational and Scientific Center of the Air Force «N.E. Zhukovsky and Y.A. Gagarin Air Force Academy», Voronezh

The article discusses the issues of improving the implementation of radio monitoring measures, taking into account the assessment of its reliability. The assessment of the priority of the directions for improving radio monitoring is considered from the perspective of solving a multi-criteria problem.

Keywords: Radio monitoring, reliability, parameter significance coefficient, effectiveness of radio monitoring.

ПОДХОД К ПОСТРОЕНИЮ МАТЕМАТИЧЕСКОЙ МОДЕЛИ ОЦЕНКИ ВОЗМОЖНОСТЕЙ СИСТЕМЫ СВЯЗИ ПО ОБЕСПЕЧЕНИЮ НЕПРЕРЫВНОСТИ УПРАВЛЕНИЯ АВИАЦИОННЫМ ФОРМИРОВАНИЕМ

*Военный учебно-научный центр Военно-воздушных сил
«Военно-воздушная академия имени профессора Н.Е. Жуковского и Ю.А. Гагарина»,
г. Воронеж*

В статье рассматривается подход к математическому описанию процесса управления авиационным формированием с позиции непрерывности, а также схематично представлена математическая модель расчета вероятности непрерывности управления.

Ключевые слова: математическая модель, непрерывность управления, система связи.

Введение

Выполнение поставленных задач силами авиационных формирований во многом зависит от качества управления ими соответствующими должностными лицами, анализ источников в данной области показал, что наиболее проработанными на сегодняшний день являются такие свойства процесса управления как устойчивость, оперативность и скрытность управления, в свою очередь наименее раскрытым, по мнению авторов, является непрерывность управления.

Под непрерывностью управления авиационным формированием в статье понимается способность системы управления обеспечить процесс управления в необходимом темпе и функционирование управляемых сил без снижения их боевой готовности, устойчивости и эффективности [1].

Любая система включает множество различных элементов, находящихся в отношениях и связях между собой и образующих определенную целостность, это положение в полной мере относится и к системе военного управления [2]. Технической основой указанной системы является система связи, которая выполняет функции по обеспечению обмена данными между элементами системы управления и степени реализации ее возможностей, во многом, определяется результат, который может быть достигнут авиационным формированием. Резюмируя сказанное, оценить возможности системы связи – это значит произвести сопоставление значений показателей, характеризующих ее возможности по обеспечению процесса управления с эталонными (нормативными) значениями, предварительно произведя измерение – алгоритмическую операцию, в результате которой получают значение показателя (количественную меру свойства), связывающего физическую интенсивность свойства с числом.

Широкое внедрение математических методов в область управления боевыми и обеспечивающими действиями авиации ставит перед собой главную цель – усовершенствовать процедуру принятия решения, сделав ее более обоснованной [3,4], рациональное применение данных методов позволяет, опираясь на результаты анализа прикладной области, построить математическую модель, и далее на ее основе дать оценку возможностей существующей системы связи авиационного формирования, а также осуществить прогноз их изменений с учетом влияния различных факторов.

Целью статьи является разработка математической модели оценки возможностей системы связи по обеспечению непрерывности управления авиационным формированием. На основе каких показателей и критериев будет построена рассматриваемая математическая модель и в чем ее отличие от существующих будет разобрано авторами в статье.

Выбор системы показателей и критериев оценки

Ресурсы системы связи служат для обеспечения обмена информацией между субъектами и объектами управления. При этом по каналу прямой связи передаются управляющие воздействия, а по каналу обратной связи – информация о состоянии объекта. Совокупность, включающую субъект управления, объекты управления и ресурсы системы связи, будем называть системой управления. Для обеспечения потребности системы управления в информационном обмене и решения задач управления на базе системы связи создаются информационные направления, которые реализуются направлениями связи. Таким образом, направления связи являются тем объектом, который связывает в единое целое систему связи (так как информационное направление базируется на направлении связи) и систему управления (так как создается в целях решения задач управления).

Наглядно влияние состояния системы связи на качество процесса управления авиационными формированиями представлено на рисунке 1.

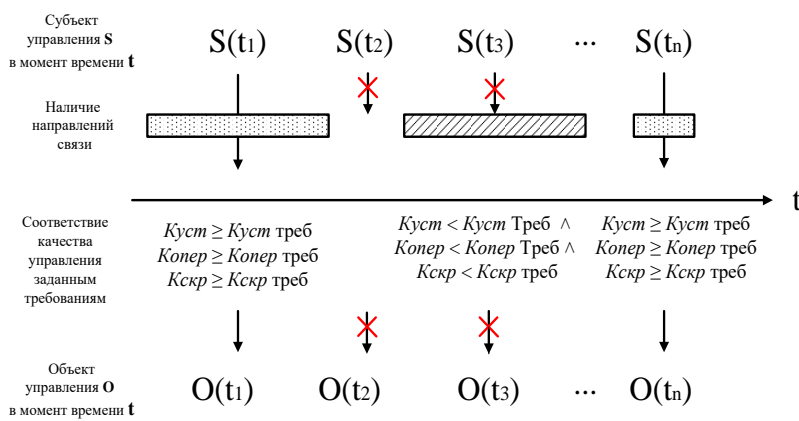


Рис. 1. Влияние состояния системы связи на качество процесса управления авиационными формированиями

Определяя критерии оценки возможностей системы связи обеспечить непрерывность прохождения управляющих воздействий авторами предлагается рассмотреть число действующих направлений связи между наземным пунктом управления и авиационным комплексом как основу для прохождения команд и отчетов в необходимом темпе, с дальнейшей оценкой оставшихся качеств управления (устойчивости, оперативности и скрытности) в рамках рассматриваемого временного интервала. На рисунке 1 управляющие воздействия t_1 и t_n удовлетворяют одновременно условиям по наличию необходимого количества направлений связи $N_{треб}$ (как характеристики непрерывности управления) и требованиям $K_{уст_треб}, K_{опер_треб}, K_{скр_треб}$ (характеризующие соответственно устойчивость, оперативность и скрытность управления). В свою очередь, в момент времени t_2 не выполняются требования по наличию необходимых направлений связи $N_{треб}$. Рассматривая систему связи в момент времени t_3 и отклонения от требований хотя бы по одному из $K_{уст_треб}, K_{опер_треб}, K_{скр_треб}$ позволяют говорить о высоковероятной потере управления авиационным комплексом.

Из чего следует, что критерием для оценки возможностей системы связи по обеспечению непрерывности управления авиационными комплексами является предоставление требуемого количества направлений связи $N_{треб}$ между бортами и наземными пунктами управления.

С позиции теории управления вероятность того, что управление авиационным

формированием будет непрерывным характеризуется вероятностью непрерывного управления, принятого за обобщенный показатель и может быть получено путем решения уравнения:

$$P_{ny}(t) = \frac{\alpha}{\alpha + \beta} [1 - e^{-(\alpha + \beta)t}] + e^{-(\alpha + \beta)t}, \quad (1)$$

$\alpha = \frac{1}{T_e}$ – интенсивность восстановления управленческих функций;

$\beta = \frac{1}{T_{np}}$ – интенсивность потери управленческих функций;

T_{np} – среднее время между интервалами потери управления;

T_e – среднее время восстановления управленческих функций.

Переходя к выбору показателей необходимо пояснить, что выполнение поставленной задачи авиационным формированием будет производиться в несколько этапов и на каждом из этапов в зависимости от складывающейся обстановки система связи должна предоставлять требуемый ресурс системы связи (направлений связи). При этом число объектов управления (авиационных комплексов) по ходу выполнения поставленной задачи может остаться постоянным (за исключением понесенных потерь), однако субъекты управления (пункты управления) могут изменяться. Это связано, прежде всего, с заданными рубежами передачи управления согласно заблаговременно созданного плана управления, так и в связи с необходимостью поддержания требуемого уровня качества управления при деструктивном воздействии противника. Данная ситуация в графическом виде показана на рисунке 2, где по оси t показано время выполнения поставленной задачи как в целом так и по этапам ее выполнения, по оси $P_{ny}(t)$ уровень непрерывности управления (при $P_{ny}(t)=1$ управление идеальное, при $P_{ny}(t)<0,8$ управление считается потерянным), а также график $U(t)_i$ характеризующие возможности системы связи i -го ПУ обеспечить процесс передачи управляющих воздействий в необходимом темпе.

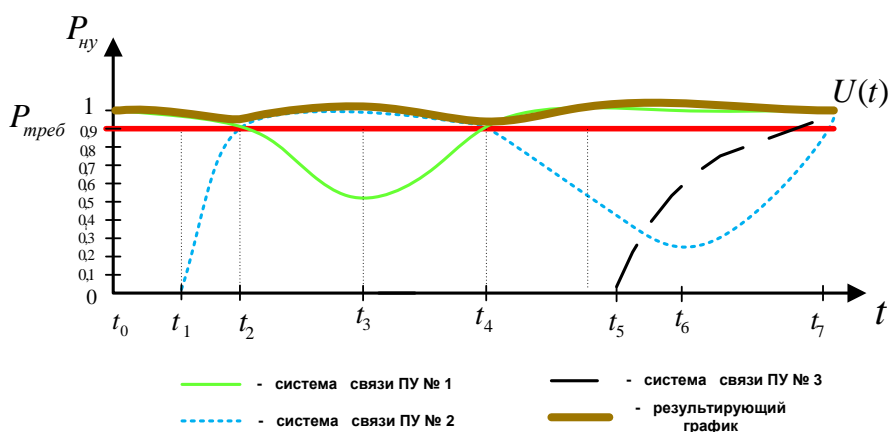


Рис. 2. График зависимости вероятности непрерывности управления от показателей T_{np} и T_e системы связи (вариант)

Приоритетным направлением при анализе возможностей системы связи i -го ПУ является установление момента времени, когда график функции $U(t)_i$ пересечет линию $P_{треб}$, что можно приравнять к потере управления объектом, и в целях недопущения указанной ситуации, система связи должна своевременно осуществить проверку состояния ресурса связи (направлений связи) на том ПУ, с которого согласно приоритета будет далее осуществляться управление, далее на основании решения должностного лица органа управления обеспечить передачу управления с одного субъекта на другой.

Если при обеспечении требуемого качества управления в исходном состоянии системы управления главенствующую роль играет показатель устойчивости системы связи, характеризующий ее способность обеспечивать управление при всех воздействующих факторах, то при решении задачи по обеспечению передачи управления с одного пункта управления на другой наиболее актуальным предлагается рассмотрение показателя мобильности ресурса системы связи P_{mob} . Указанный показатель, характеризует возможность оперативного переключения каналов связи на приоритетный субъект управления, а также использование полевой компоненты системы связи для наращивания своих возможностей на избранном информационном направлении. В комплексе данное решение позволит с момента времени t_2 обеспечить функционирование необходимого числа направлений связи и своевременно прекратить снижение вероятности непрерывного управления, а также предоставит время для восстановления состояния системы связи до требуемого уровня.

Также среди показателей характеризующих возможности системы связи по обеспечению непрерывности управления авиационными формированиями предлагается рассмотреть доступность ресурса системы связи $P_{дост}$, позволяющий учесть обеспечение возможности должностным лицам органов и пунктов управления различных звеньев получения доступа к ресурсам системы связи при сохранении установленных приоритетов и способов установления связи.

На основании вышеизложенного, при оценке возможностей системы связи по обеспечению непрерывности управления авиационными формированиями может быть применена следующая система показателей:

обобщенный показатель: вероятность непрерывности управления авиационным формированием - $P_{ny}(t)$;

общие показатели: вероятность сохранения требуемого количества функционирующих направлений связи $P_{R_соxp}$, а также вероятность восстановления требуемого количества функционирующих направлений связи $P_{R_восст}$;

частные показатели: вероятность устойчивости требуемого ресурса системы связи $P_{уст}$, вероятность мобильности требуемого ресурса системы связи P_{mob} , вероятность доступности требуемого ресурса системы связи $P_{дост}$.

Построение математической модели оценки возможностей системы связи по обеспечению непрерывности управления авиационным формированием

Приведенные рассуждения можно представить в виде математической модели, которая в общем виде позволит описать порядок проведения расчетом искомой величины – вероятности непрерывности управления авиационным формированием при выполнении поставленной задачи. Очевидным преимуществом данной математической модели является учет деструктивного воздействия противника, а также динамики выполнения поставленной задачи за счет применения указанных исходных данных при расчетах возможностей системы связи и учета важности направлений связи при определении вероятности ее устойчивости.

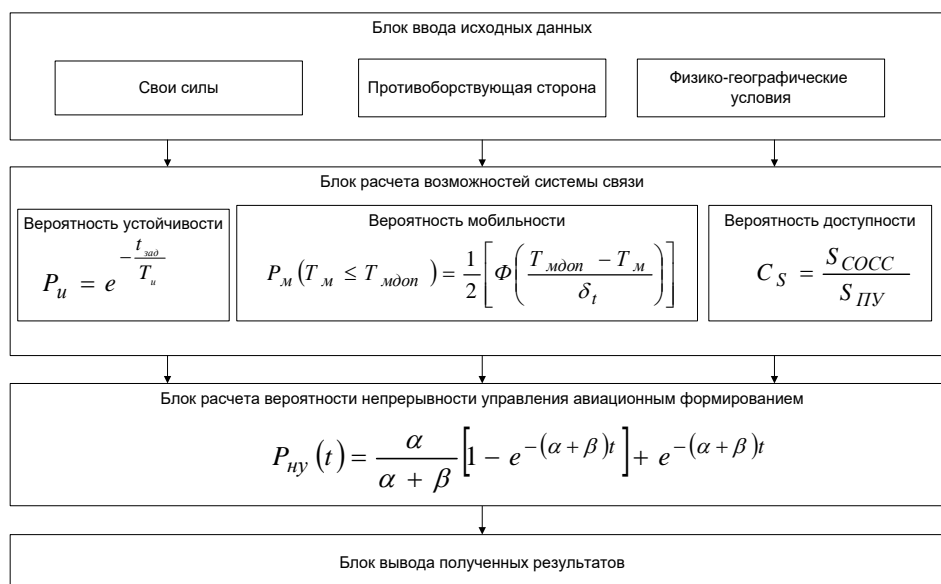


Рис. 3. Математическая модель оценки возможностей системы связи по обеспечению непрерывности управления авиационным формированием (вариант)

Заключение

Математическое моделирование при исследовании выполнения поставленных задач авиационными формированиями оказывается практически единственным средством познания вследствие трудностей постановки экспериментов и наблюдения явлений в реальных условиях, оно позволяет не только учитывать динамику взаимодействия элементов и параметров систем в пространстве и во времени, но и предоставляет возможность изменять масштаб времени и описывать поведение компонентов сложной системы на высоком уровне детализации.

Однако необходимо упомянуть, что какой хорошей бы ни была созданная модель изучаемой системы, основной интерес для командиров всех звеньев управления представляет не сама математическая модель, а получаемые с ее помощью и интерпретируемые результаты решения реальной проблемы. Кроме того, информация, получаемая с помощью модели, должна быть приемлемой с практической точки зрения, тогда можно говорить о том, что получен мощный и гибкий «инструмент» для получения научных результатов в области исследования.

Библиографический список

1. Основы теории управления войсками / [П. К. Алтухов, И. А. Афонский, И. В. Рыболовский, А. Е. Татарченко; Под ред. П. К. Алтухова. - М.: Воениздат, 1984. -221 с.
2. Основы и применение методов прикладной математики в военном деле/ Иванов П.И., учебник – Монино, 1991. - 512 с.
3. Понятийный аппарат в предметных областях познания и исследования военного искусства (категории, понятия, термины): учебно-наглядное пособие / Махнин В.Л., Бычков В.Г. - М.: Изд-во ВУНЦ ВВС «ВВА им. проф. Н.Е. Жуковского и Ю.А. Гагарина», 2011. – 324 с.
4. Основы военно-научных исследований: понятийный аппарат: монография/ Махнин В.Л.- Воронеж: ВУНЦ ВВС «ВВА», 2017. – 460 с.

S.S. Belousov, P A. Fedyunin, M.A. Stafeev

**AN APPROACH TO THE CONSTRUCTION OF A MATHEMATICAL MODEL
FOR ASSESSING THE CAPABILITIES OF A COMMUNICATION SYSTEM
TO ENSURE THE CONTINUITY OF AVIATION FORMATION MANAGEMENT**

*Military educational and scientific center of the Air Force Military
«Air Force Academy», Voronezh*

In article considers an approach to the mathematical description of the process of controlling an aviation formation from the point of view of continuity, and also schematically presents a mathematical model for calculating the probability of continuity of control.

Keywords: mathematical model; the continuity of the control system.

МОДЕЛЬ ОЦЕНКИ РАДИОТЕХНИЧЕСКОГО ОБЕСПЕЧЕНИЯ ГОСУДАРСТВЕННОЙ АВИАЦИИ В УСЛОВИЯХ ДЕСТРУКТИВНОГО ВОЗДЕЙСТВИЯ ПРОТИВНИКА

*Военный учебно-научный центр Военно-воздушных сил
«Военно-воздушная академия имени профессора Н.Е. Жуковского и Ю.А. Гагарина»,
г. Воронеж*

В статье представлена модель оценки радиотехнического обеспечения государственной авиации в условиях разрушительного воздействия противника и предложен подход к выбору показателей оценки эффективности радиотехнического обеспечения. Рассматриваемая модель может быть использована в качестве основы для рекомендаций по совершенствованию радиотехнического обеспечения государственной авиации.

Ключевые слова: модель, радиотехническое обеспечение, государственная авиация, эффективность.

На современном этапе развития Вооруженных Сил, с учетом роста решаемых задач в информационной сфере, происходит стремительное развитие информационных, телекоммуникационных технологий, появляются новые способы передачи информации, а также средства и способы противодействия ей. Такие изменения требуют новой оценки процессов обеспечения управления государственной авиации в воздухе в условиях деструктивного воздействия противника и определение направлений развития и способов организации радиотехнического обеспечения (РТО) полетов. Кроме того, анализ существующих подходов к оценке данного вида обеспечения полетов не учитывают специфику боевого применения государственной авиации, а также возможности поступающих на вооружение современных средств РТО.

Таким образом, возникает необходимость совершенствования научно-методического аппарата в части касающейся оценки возможностей РТО полетов государственной авиации в современных условиях ведения военных действий.

Выполнение задач государственной авиации осуществляется в ходе выполнения полетов как по определенным воздушным трассам, так и вне воздушных трасс, при этом каждый полет воздушного судна условно можно разбить на следующие этапы: взлет и построение боевого порядка; полет в район выполнения авиационных работ; действия в районе выполнения авиационных работ; полет в район аэродромов посадки; роспуск боевых порядков и выполнение посадки. Таким образом, справедливо предположить, что эффективность радиотехнического обеспечения государственной авиации определяется успешностью выполнения каждого этапа полета воздушного судна.

Каждому этапу полета свойственны свои цели и содержание, исходя из которых, определяются цели РТО, заключающиеся в создании экипажам самолетов благоприятных условий для успешного выполнения полетного задания.

Процессы взлета и построения боевых порядков влияют на выдерживание установленной последовательности действий экипажей на последующих этапах. Взлет воздушного судна осуществляется в установленное время, по сигналу и по вызову из различных степеней готовности воздушного судна. При взлете в установленное время запуск осуществляется с учетом потребного для выруливания на взлетно-посадочную полосу времени.

Общими требованиями к взлету и построению боевых порядков являются обеспечение своевременности, скрытности, безопасности и максимального продвижения

по маршруту. При выполнении взлета экипажам при рулении и разбеге необходимо выдерживать направление, ориентируясь по огням магистральной рулежной дорожки, взлетно-посадочной полосы и взлетным огням, при этом дальность видимости огней светосигнального оборудования должна составлять не менее требуемой, при этом учитывается максимальная скорость движения воздушного судна по взлетно-посадочной полосе. Основная задача РТО полетов на данном этапе заключается в своевременном обеспечении летчику ночью или в сложных метеоусловиях надежного визуального контакта с огнями светосигнального оборудования системы посадки аэродрома в целях точного выдерживания направления взлета относительно осевой линии или боковых сторон взлетно-посадочной полосы.

Выдерживание заданного места в боевом порядке производится экипажами визуально. При действиях в условиях, исключающих возможность визуального контроля, оно должно осуществляться с помощью радиотехнических средств. Для контроля за построением боевых порядков, предупреждения экипажей об опасных сближениях и недопущения столкновений группа руководства полетами будет использовать информацию от радиолокационных средств. Особенно это важно для воздушных судов, которые не имеют аппаратуру межсамолетной навигации.

После взлета, сбора и построения боевого порядка начинается полет по заданному маршруту с целью своевременного, точного в район выполнения авиационных работ. На этом этапе полета РТО решает задачи формирования и выдачи информации с такой точностью и дискретностью, которая позволит экипажу выдерживать заданный маршрут полета, отыскать назначенный район, выйти на него и выполнить поставленную задачу. Успешность решения данной задачи определяется точностью выдерживания линии заданного пути на всем протяжении маршрута.

Особенность обеспечения данного этапа полета является то, что перед РТО стоят задачи по обеспечению точного вывода экипажей в область работы прицельно-навигационного комплекса воздушного судна в простых и сложных метеоусловиях, после чего экипажи выполняют самостоятельный поиск назначенного ориентира визуально.

Используя параметры прицельных систем воздушного судна можно предъявить минимальные требования к радиотехническому обеспечению вывода самолетов на наземные объекты, а именно в ту область воздушного пространства, в которой будет осуществляться гарантированное применение воздушных судов государственной авиации. Следовательно, задача РТО при обеспечении данного этапа полета будет заключаться в выводе экипажей в область воздушного пространства со следующими характеристиками: минимальная и максимальная дальность до объекта при высоте, обеспечивающей прямую его видимость. Дальность прямой видимости определяется исходя из метеорологической обстановки в районе выполнения авиационных работ. В статье принято, что наихудшими условиями для их выполнения являются сложные метеорологические условия [1].

После выполнения авиационных работ экипажи воздушных судов возвращаются на аэродромы. Обратный маршрут, как правило, может не совпадать с маршрутом полета к району выполнения авиационных работ и осуществляется в назначенных полосах пролета. Задача РТО при этом аналогична второму этапу.

С выходом в район аэродрома будет осуществляться роспуск боевого порядка, заход на посадку и посадка. Цель этапа – выполнение успешной посадки с первого захода, исключающей потери летного состава и авиационной техники. На заключительном этапе полета перед РТО стоят задачи по обеспечению вывода экипажей в точку начала снижения и информацией о местоположении самолета в воздушном пространстве относительно аэродрома и на посадочном курсе, об отклонениях от линии западного пути, по обеспечению визуального контакта с взлетно-посадочной полосой и о местах стоянок. Значения точности выхода в точку начала снижения зависят от точностей: определения путевой скорости, ее выдерживания, выдерживания курса полета [2].

Таким образом, стало возможным разработать алгоритм и модель оценки РТО государственной авиации в условиях деструктивного воздействия противника, представленные на рисунках 1 и 2 соответственно, которые позволили в дальнейшем разработать программный продукт в помощь должностным лицам, непосредственно осуществляющим организацию РТО полетов государственной авиации.

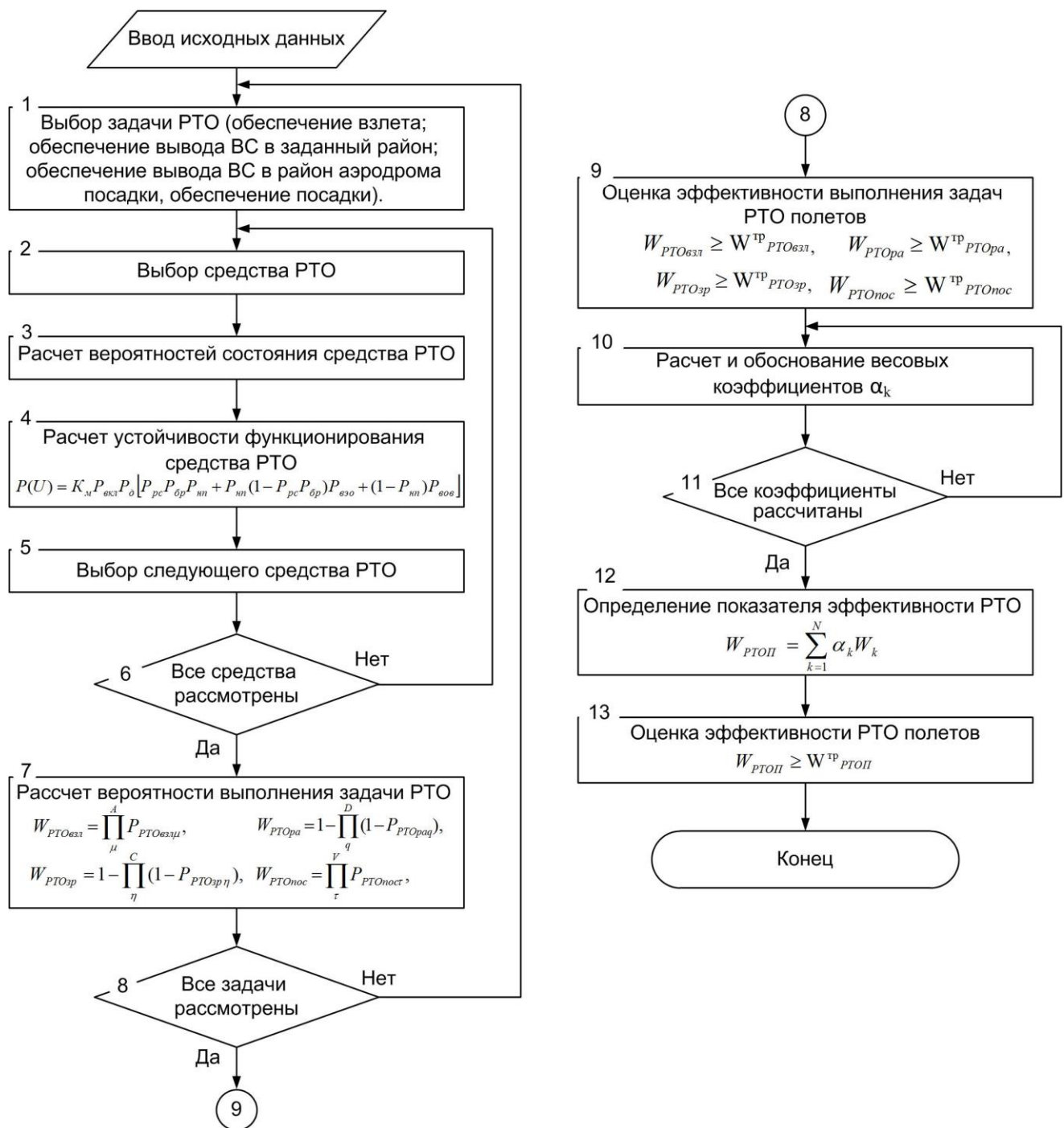


Рис. 1. Алгоритм оценки РТО государственной авиации в условиях деструктивного воздействия противника

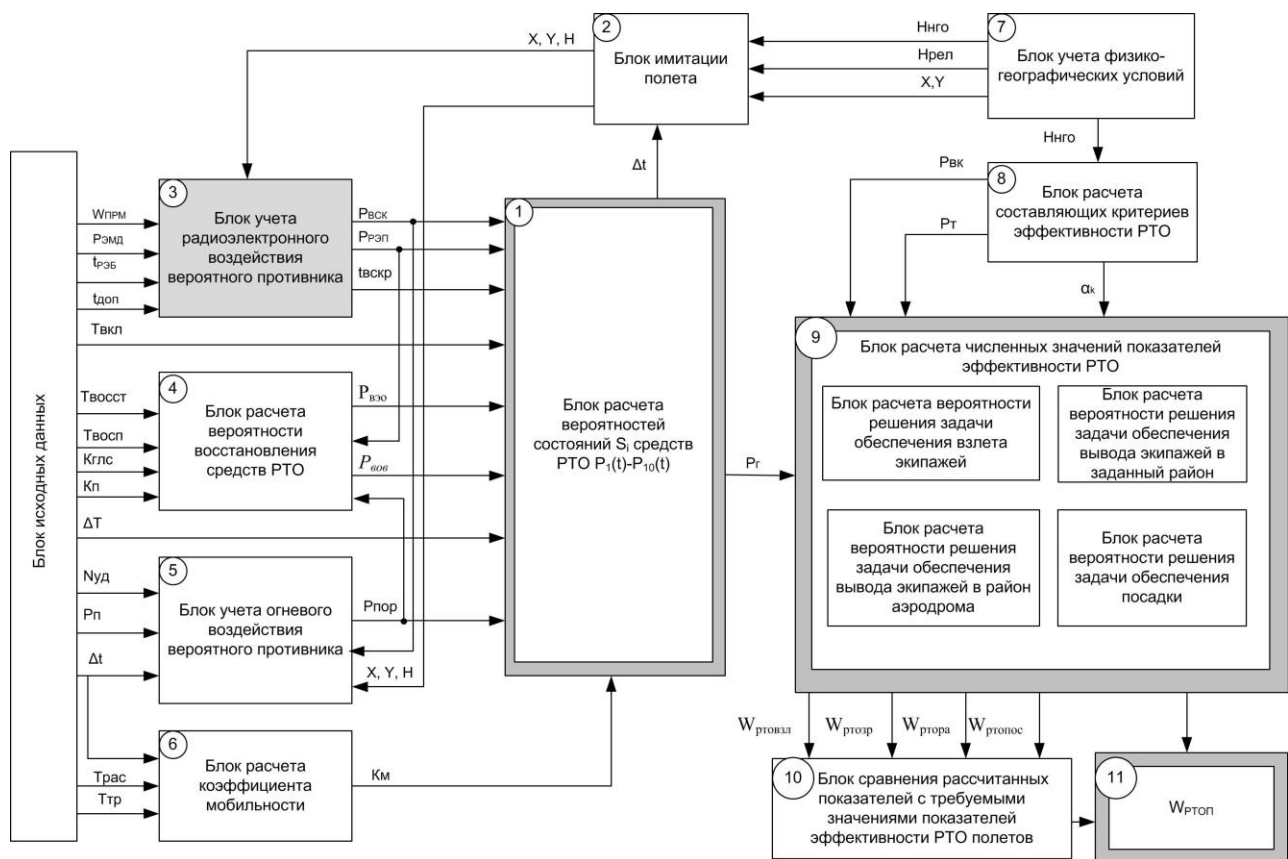


Рис. 2. Методика оценки государственной авиации в условиях деструктивного воздействия противника

Для осуществления вычислений был разработан программный продукт с использованием табличного процессора MSExcel и встроенного языка программирования VBA. В основе вычислительного ядра лежит метод решения систем линейных уравнений методом обратной матрицы. Программный продукт позволяет осуществлять расчет вероятностей состояний P_i , а также итоговой вероятности $P_{РТО}$ для разных условий функционирования одиночного средства РТО и действий противника.

Таким образом, процесс управления полетами государственной авиации характеризуется иерархичностью организации, целенаправленностью функционирования, большим числом объектов управления, наличием множества информационных связей и взаимодействия между ними. Для реализации функций управления полетами государственной авиации создается система управления, технической основой которой является система радиотехнического обеспечения. Из рассмотренных целей и содержания этапов полета воздушного судна государственной авиации следует, что на радиотехническое обеспечение полетов возлагается решение следующих основных задач: обеспечение взлета экипажей с основного аэродрома базирования; обеспечение вывода экипажей воздушных судов в район выполнения авиационных работ; обеспечение вывода экипажей воздушных судов в район аэродрома посадки; обеспечение посадки.

Библиографический список

1. Ивануткин А.Г. Методика оценки эффективности радиотехнического обеспечения полётов авиации // Военная мысль. Выпуск №7. – 2016. С.33-40.
2. Горобец Г.С. Теория вероятностей, математическая статистика и элементы случайных процессов / Г.С. Горобец. - М.: КД Либроком, 2013. - 232 с.

A.G. Ivanutkin, A.S. Demyanenko, A.V. Blinov

MODEL FOR EVALUATION OF RADIO-TECHNICAL SUPPORT OF STATE AVIATION IN CONDITIONS OF DESTRUCTIVE IMPACT OF THE ENEMY

*Military educational and scientific center of the Air Force Military «Air Force Academy»,
Voronezh*

The article presents model for evaluation of radio-technical support of state aviation in conditions of destructive impact of the enemy and an approach to the selection of indicators to assess the effectiveness of radio-technical support. The considered model can be used as a basis for recommendations on improving of radio-technical support of state aviation.

Keywords: model, radio-technical support, state aviation, efficiency.

СОВЕРШЕНСТВОВАНИЕ РАДИОСВЯЗИ В АВИАЦИОННЫХ ФОРМИРОВАНИЯХ НА ОСНОВЕ ПРИМЕНЕНИЯ ПРОГРАММНО- ОРИЕНТИРОВАННОГО РАДИО

*Военный учебно-научный центр Военно-воздушных сил
«Военно-воздушная академия имени профессора Н.Е. Жуковского и Ю.А. Гагарина»,
г. Воронеж*

В статье представлено исследование радиосвязи декаметрового диапазона в рамках системы связи авиационного формирования в прогнозируемых условиях радиоэлектронного противодействия противника. В качестве основного направления совершенствования радиосвязи авиационных формирований предложено применение программно-ориентированного радио.

Ключевые слова: радиосвязь, радиоэлектронное подавление, программно-ориентированное радио.

На практике радиосвязь авиационных формирований в декаметровом диапазоне длин волн (ДКМВ) носит проблемный характер, это связано с дефицитом частотного ресурса, технической готовностью радиосредств и уровнем подготовки операторов, а также с радиоэлектронным воздействием противника. Анализ факторов, влияющих на информационный обмен в радиолиниях ДКМВ диапазона, показывает, что наиболее существенное влияние на радиосвязь оказывают: трудно прогнозируемое состояние ионосферы (в краткосрочной перспективе) и деструктивное воздействие противника (огневое и функциональное поражение, радиоэлектронное подавление и др.) [1, 2].

Силы и средства радиоэлектронного подавления (РЭП) противника постоянно совершенствуются, принимаются на вооружение роботизированные комплексы, в качестве помех сложные радиосигналы и даже искусственный интеллект. Из открытых источников [2] и информационно-телекоммуникационной сети «интернет» известно, что технически развитые государства исследуют воздействие на ионосферу с целью изменения условий распространения радиоволн, это может быть «накачка» ионосферы мощным сверхвысокочастотным излучением, создание искусственных неоднородностей (зеркал) и др. Направлениями исследований являются создание благоприятных условий для обеспечения своей радиосвязи в определенных районах и затруднения информационного обмена противоборствующей стороне.

В таких сложных условиях в авиационных формированиях сокращаются штатные силы и средства частотно-диспетчерской службы, нарушаются устоявшиеся принципы частотного планирования, такие как разделение участков частотного спектра между видами и родами войск. Вместе с тем наблюдается экспоненциальный рост численности объектов управления и радиосредств для управления ими, применение методов расширения базы радиосигналов (программная перестройка рабочей частоты, широкополосные сигналы и др.) и снижение качества планирования радиосвязи во всех органах военного управления из-за высокого уровня неопределенности. Вышеперечисленное не гарантирует в авиационных формированиях качество радиосвязи при выполнении задач в вооруженных конфликтах.

В прогнозируемых условиях радиоэлектронного противодействия противника осуществление бесподстроечной радиосвязи на фиксированных частотах в ДКМВ диапазоне между пунктами управления (ПУ), ПУ и экипажами воздушных судов, по мнению авторов, представляется маловероятным.

Таким образом, подбор рабочей частоты или группы частот в ДКМВ диапазоне будет занимать значительное время, которое надо учитывать при организации радиосвязи.

$$T_{ур/св} = T_{пч} + T_{настр} + T_{вх р/св}$$

где $T_{ур/св}$ – время установления радиосвязи; $T_{пч}$ – время подбора рабочей частоты (группы); $T_{настр}$ – время настройки радиостанции; $T_{вх р/св}$ – время вхождения в связь с корреспондентом.

Элементная база современных радиостанций позволяет сокращать время настройки и перестройки рабочих частот, а также автоматически входить в связь (путем отслеживания работы на несущей частоте, обмена служебными командами и др.), в тоже время разработке техники частотно-диспетчерской службы уделяется меньше внимания. Если ранее на узлах связи авиационных формирований применялись специализированная аппаратура, обладающая возможностями вертикального и наклонного зондирования ионосферы, то в настоящее время отсутствуют штатные должностные лица и техника, позволяющие в полном объеме осуществлять работу по подбору рабочих частот и кратковременному прогнозированию их использования.

Одним из направлений совершенствования радиосвязи авиационных формирований предлагается применение Software-Defined Radio (SDR) в отечественной литературе чаще применяется понятие программно-ориентированное радио (ПОР) или программно-конфигурируемое радио. ПОР реализует передачу, прием и обработку сигналов в программном виде на нулевой либо промежуточной частоте с последующим аппаратным переносом на промежуточную радиочастоту. Таким образом, ПОР является радиооборудованием, в котором все или большинство функций физического уровня выполняются в программном виде [3]. В настоящее время развитие аналого-цифровых и цифро-аналоговых преобразователей, вычислительных мощностей процессоров (ЭВМ) обеспечивают возможность программной обработки сигналов непосредственно на рабочей частоте. Совместное применение ПОР радиоприемника с ЭВМ позволяет визуализировать на мониторе широкий участок радио спектра в режиме времени близком к реальному, обнаруживать излучение радиосигналов, их параметры, выполнять фильтрацию, выбирать вид модуляции без внесения изменений в аппаратную часть и другие преимущества.

Устройства на базе ПОР технологий позволяют в разы сократить продолжительность разработки прогноза распространения радиоволн при вертикальном и наклонном зондировании ионосферы, а также открывает возможности обработки (применения) линейных частотно модулированных сигналов.

В прогнозируемых условиях вооруженного противоборства РЭП противник будет осуществлять с применением помех сложной сигнальной конструкции, интеллектуальное подавление (при наличии излучения радиосигнала), временное уплотнение работы передатчиков помех (подавление на нескольких частотах с определенной скважностью), передачей сигналов и команд, записанных ранее, их синтезированные вариации. На практике радиотелефонисту (телеграфисту) будет почти невозможно по встроенным приборам радиоприемника и на слух определить наличие преднамеренных помех и вообще осуществляется РЭП противником или имеется неисправность аппаратуры. ПОР приемник со средствами индикации позволяет визуально определить наличие преднамеренных помех на рабочих частотах и соответственно принимать меры по радиоэлектронной защите радиолиний.

Целесообразность графического представления принимаемого сигнала в ДКМВ диапазоне обусловлена наличием помех преднамеренных и непреднамеренных помех, многолучевостью распространения радиоволн, что приводит к снижению вероятности приема сообщений в автоматизированных радиосистемах. В отличие от слухового восприятия человека, способного адаптироваться к посторонним шумам и определять наличие полезного сигнала даже при соотношении сигнал/шум менее 1, зрительное восприятие заранее заданного образа (геометрической фигуры, отрезка, линии и т.д.) обеспечивается при потере или искажении до 90 % элементов.

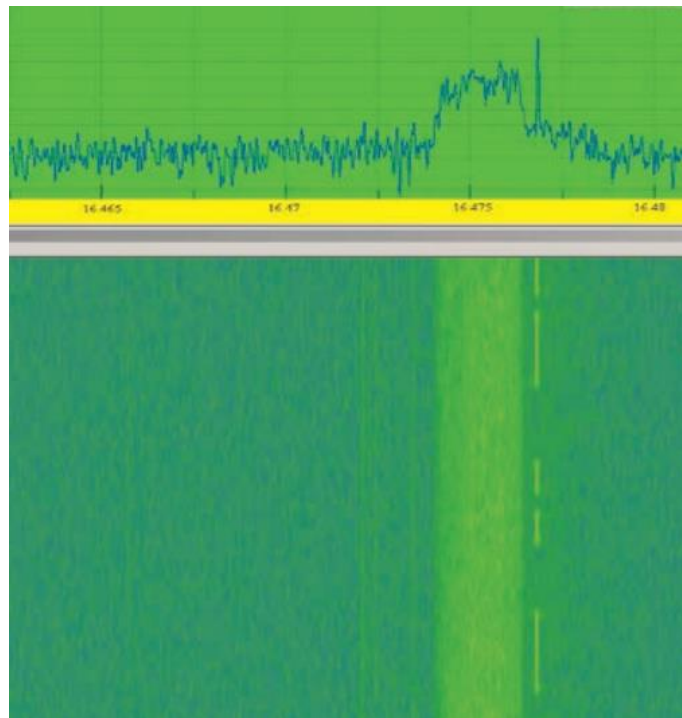


Рис. 1. Визуальный прием азбуки Морзе

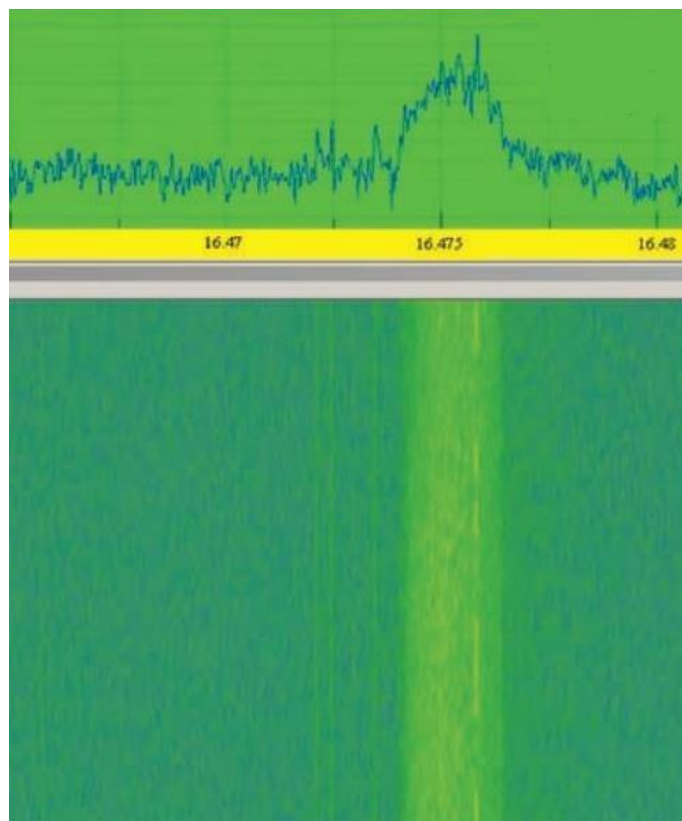


Рис. 2. Визуальный прием азбуки Морзе на фоне радиопомехи

Это определяется важной особенностью нейронной сети мозга человека по восстановлению полного образа, хранящегося в памяти, по неполным и искаженным данным [4]. Таким образом, визуальный прием сигналов азбуки Морзе, с использованием приемника, отображающего на мониторе временную спектрограмму сигнала, позволяет зрительно зарегистрировать передаваемую информацию, даже в том случае, когда в

слуховом режиме прием невозможен из-за низкого соотношения сигнал/шум (помеха) на рисунках 1 и 2.

Таким образом, в современных условиях технически развитого вооруженного противоборства обеспечение качественной радиосвязи в ДКМВ диапазоне невозможно без организации частотно-диспетчерской службы на узлах связи авиационных формирований. Одним из направлений совершенствования которой является применение программно-ориентированного радио, что позволяет визуально оценивать занятость частот в широкой полосе, оперативно контролировать работу своих радиостанций, прогнозировать радиоэлектронную обстановку в районе узлов связи, выявлять преднамеренные помехи. При организации радиосвязи в авиационном формировании необходимо учитывать время на подбор рабочих частот, а также время перестройки на другие частоты, так как вероятность занятости и радиоэлектронного противодействия противником очень высока.

Библиографический список

1. Связь в Вооруженных силах Российской Федерации - 2020. - М.: «Информационный мост», 2020. - 250 с.
2. Сидоркин А.Н. Информационные, специальные, воздушно-десантные и аэромобильные операции армий ведущих зарубежных государств: Информационно-аналитический сборник / А.Н. Сидоркин, И.А. Рябенко, В.П. Герасимов. - М.: Воениздат, 2011. - 344 с.
3. Фокин Г.А. Технологии программно-конфигурируемого радио / Г.А. Фокин. - М.: Горячая линия - Телеком, 2019. - 316с.
4. Николашин Ю.Л. Эффективность использования когнитивной радиосвязи в декаметровом диапазоне частот / Ю.Л. Николашин, П.А. Будко, Г.А. Жуков // Техника средств связи № 2 (142). - СПб.: ПАО «Интелтех», 2018. - С.6-20.

M.A. Danilin, A.N. Tatarinov

IMPROVEMENT OF RADIO COMMUNICATION IN AVIATION UNITS BASED ON THE USE OF SOFTWARE-DEFINED RADIO

*Military educational and scientific center of the Air Force Military «Air Force Academy»
(Voronezh)*

The article presents a study of decameter radio communication in framework of communication system of aviation units in predicted conditions of enemy's electronic counteraction. As the main direction of improving radio communication of aviation units, use of software-defined radio is proposed.

Keywords: radio, electronic suppression, software-defined radio.

ОЦЕНКА ЭФФЕКТИВНОСТИ СИСТЕМЫ ТЕХНИЧЕСКОГО ОБСЛУЖИВАНИЯ И РЕМОНТА РАДИОТЕХНИЧЕСКИХ УСТРОЙСТВ

*Военный учебно-научный центр Военно-воздушных сил
«Военно-воздушная академия имени профессора Н.Е. Жуковского и Ю.А. Гагарина»,
г. Воронеж*

В статье рассмотрен подход к оценке эффективности системы технического обслуживания и ремонта, которая позволяет функционально увязать параметры системы с требуемым составом радиотехнических устройств, нормами их поставки и затратами на их эксплуатацию. В качестве математического аппарата выбрана система дифференциальных уравнений Колмогорова.

Ключевые слова: радиотехническое устройство, система технического обслуживания и ремонта, система дифференциальных уравнений Колмогорова, эксплуатация, ремонт, стоимость.

Введение

При помощи специальных математических моделей можно произвести научное прогнозирование состояния системы технического обслуживания и ремонта (ТОиР) и определить ее эффективность. В соответствии с вышеизложенным [1] оценку эффективности системы ТОиР целесообразно проводить в три этапа:

1. Обоснование требуемой рациональной системы ТОиР, при которой обеспечивается требуемый уровень эффективности W_{mp} при минимальных затратах C_{min} .
2. Определение уровня эффективности реальной системы ТОиР (W, C).
3. Сравнение показателей эффективности (W, C) реальной системы с показателями эффективности (W_{mp}, C_{min}) требуемой системы.

По результатам сравнения могут обосновываться предложения по совершенствованию системы ТОиР. Рассмотрим возможный подход к решению поставленных задач

Основная часть

Задача обоснования рациональной системы ТОиР состоит в определении таких эксплуатационных характеристик радиотехнических устройств (РТУ), параметров систем технического обслуживания, материально-технического снабжения, структуры и состава войсковых ремонтных органов, резервов РТУ, комплектов запасных частей, инструментов и принадлежностей (ЗИП), средств их доставки, при которых в установленные сроки достигается требуемый уровень надсистемы [1]

$$W[S(z)] \geq W_{mp}, \quad (1)$$

при минимальных стоимостных затратах

$$\{C_{PTC} + C_{PO} + C_{TO} + C_{ЗИП} + C_{рез} + C_{достав}\} \rightarrow \min, \quad (2)$$

где z – множество РТУ, характеризующихся $S(z)$ – заданной структурой сложной технической системы, в виде которой представлена система ТОиР.

Таким образом, задача определения требуемых значений параметров системы ТОиР формулируется следующим образом:

требуется определить такие значения параметров системы Z , при которых

обеспечивается наличие исправных изделий РТУ данного типа в системе в количестве m_r , не менее требуемого m_o , при продолжительности жизненного цикла $t_{жс}$ не превышающем время морального старения РТУ $t_{мс}$, и минимальных затратах C на производство и эксплуатацию изделий РТУ данного типа в составе системы.

Для решения сформулированной задачи рассмотрим множество РТУ РЛ-системы как некоторую динамическую систему, состоящую из достаточно большого количества m образцов вооружения в момент времени t .

В процессе функционирования такой системы каждый образец РТУ под воздействием различных причин может последовательно переходить из одного состояния в другое. Анализ процесса эксплуатации показал, что на стадии создания и эксплуатации образцов РТУ целесообразно выделить следующие основные состояния:

$S1$ – образец РТУ исправен, применяется (может применяться) по назначению;

$S2$ – неисправен, проводится текущий ремонт;

$S3, S4, S5$ – находиться на техническом обслуживании ТО-1, ТО-2, СО соответственно;

$S6$ – снят с эксплуатации, проводится средний (профилактический) ремонт;

$S7$ – снят с эксплуатации, проводится капитальный ремонт;

$S8$ – разворачивается после поступления из промышленности или из капитального ремонта;

$S9$ – подготовлен заводом-изготовителем к отправке к месту эксплуатации;

$S10$ – выработал ресурс, снят с эксплуатации и подлежит списанию.

Процесс смены этих состояний при определенных допущениях может быть представлен транзитивным однородным марковским с конечным числом состояний $S(t) = \{S1, S2, \dots, Sn\}$ и непрерывным временем. В данном случае $n = 9$. Размеченный граф состояний каждого образца системы представлен на рисунке 1, где индексами λ_{ij} обозначены интенсивности переходов образца РТУ из i -го возможного состояния в j -тое [2].

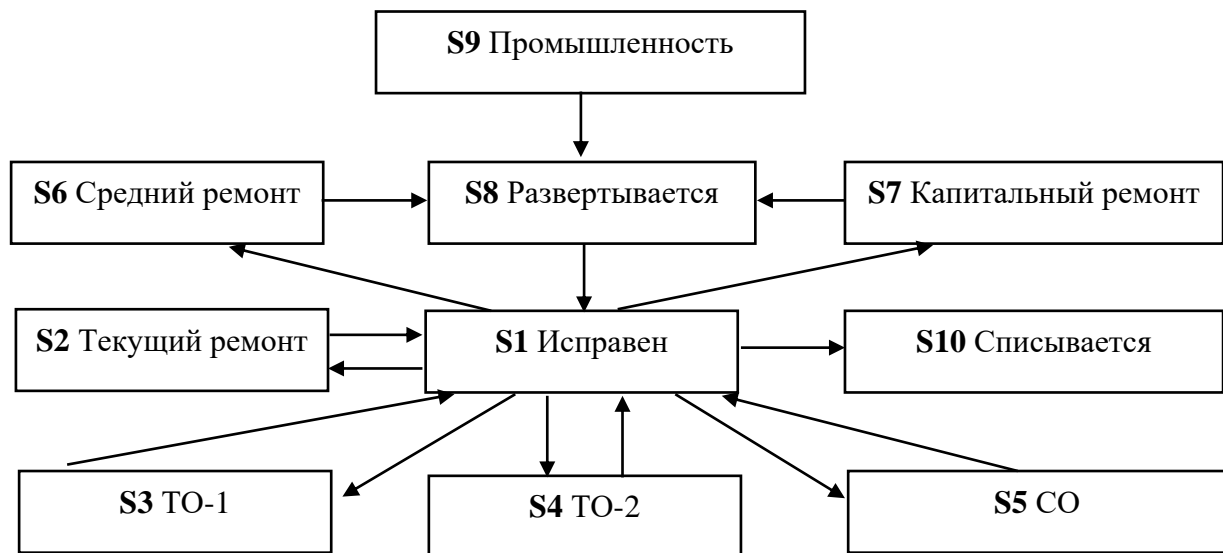


Рис. 1. Размеченный граф состояний РТУ

Поскольку потоки обладают марковским свойством, то они пуассоновские и при этом время пребывания образцов РТУ в каждом состоянии распределено по экспоненциальному закону, а интенсивности переходов λ_{ij}^* численно равны величинам, обратным среднему времени пребывания в i -том состоянии. Например, интенсивность поступления в состояние технического обслуживания ТО-1 вычисляется по формуле

$$\lambda_{3,1} = \frac{1}{T_{TO-1}}, \quad (3)$$

где T_{TO-1} – периодичность проведения технического обслуживания ТО-1.

Аналогичным образом могут быть определены и другие интенсивности переходов РТУ. Состояние системы в целом в каждый момент времени t характеризуется численностью состояний $m_k(t)$, т.е. числом образцов РТУ, находящихся в состоянии S_n , которое будет случайным и зависящим от λ_{ij} .

Общее количество образцов РТУ будет изменяться во времени, а нормирующее условие будет иметь вид:

$$\sum_{i=1}^I m_i = m(t),$$

$$m(t) = m(0) + \int_0^t \delta(t) dt - \int_0^t \lambda_{1,10} m(t) dt, \quad (4)$$

где $m(t)$, $m(0)$ – общее количество образцов РТУ системы в произвольный момент времени t и $t=0$ соответственно.

Составим систему дифференциальных уравнений Колмогорова:

$$\frac{dm1}{dt} = m2\lambda_{21} + m3\lambda_{31} + m4\lambda_{41} + m5\lambda_{51} + m8\lambda_{81} - (\lambda_{12} + \lambda_{13} + \lambda_{14} + \lambda_{15} + \lambda_{16} + \lambda_{17} + \lambda_{18})m1;$$

$$\frac{dm2}{dt} = m1\lambda_{12} - m2\lambda_{21};$$

$$\frac{dm3}{dt} = m1\lambda_{13} - m3\lambda_{31};$$

$$\frac{dm4}{dt} = m1\lambda_{14} - m4\lambda_{41};$$

$$\frac{dm5}{dt} = m1\lambda_{15} - m5\lambda_{51};$$

$$\frac{dm6}{dt} = m1\lambda_{16} - m6\lambda_{68};$$

$$\frac{dm7}{dt} = m1\lambda_{17} - m7\lambda_{78};$$

$$\frac{dm8}{dt} = m6\lambda_{68} + m7\lambda_{78} - m8\lambda_{81}.$$

Можно функционально связать выше упомянутые величины (m_{mp} , T_{TO-1} и др.), а требуемые значения этих величин определить в результате решения полученной системы дифференциальных уравнений. Систему уравнений следует решать при начальных условиях:

$$\begin{aligned}
m1(t=0) &= m10; \\
m2(t=0) &= m20; \\
\dots\dots\dots; \\
m9(t=0) &= m90.
\end{aligned}
\tag{6}$$

С учетом нормирующего условия и требования по количеству исправных образцов РТУ $m1(t) > m_{mp} = m_0$. Для установившегося процесса эксплуатации РТУ, при котором естественная убыль техники за счет ее списания компенсируется поставками новой техники из промышленности, можно использовать метод динамики средних [2]. Для этого правые части системы уравнений (5) необходимо приравнять к нулю:

$$\begin{aligned}
& m2\lambda_{21} + m3\lambda_{31} + m4\lambda_{41} + m5\lambda_{51} + m8\lambda_{81} - \\
& - (\lambda_{12} + \lambda_{13} + \lambda_{14} + \lambda_{15} + \lambda_{16} + \lambda_{17} + \lambda_{18})m1 = 0; \\
& m1\lambda_{12} - m2\lambda_{21} = 0; \\
& m1\lambda_{13} - m3\lambda_{31} = 0; \\
& m1\lambda_{14} - m4\lambda_{41} = 0; \\
& m1\lambda_{15} - m5\lambda_{51} = 0; \\
& m1\lambda_{16} - m6\lambda_{68} = 0; \\
& m1\lambda_{17} - m7\lambda_{78} = 0; \\
& m6\lambda_{68} + m7\lambda_{78} - m8\lambda_{81} = 0.
\end{aligned}
\tag{7}$$

При определенных условиях возможно получение множества допустимых значений параметров системы Z , при которых выполняется условие $m1 > m_{mp}$. В этом случае из множества допустимых значений Z выбираются такие, при которых обеспечиваются минимальные затраты на производство C_n и эксплуатацию $C_э$ образцов РТУ:

$$C_n + C_э \rightarrow \min \tag{8}$$

При определенных допущениях стоимость производства образца можно представить функционалом:

$$C_n = C_{na} \prod_i^{\beta_i} \frac{T_i}{T_{i0}}, \tag{9}$$

где C_{na} – стоимость аналога образца РТУ; T_i – i -ая эксплуатационная характеристика образца РТУ, T_{i0} – i -ая эксплуатационная характеристика аналога образца РТУ; β_i – коэффициент, характеризующий нелинейность зависимости $C = C(T_i, T_{i0})$.

Стоимость эксплуатационных затрат $C_э$ определяется в соответствии с действующей в настоящее время методикой [3] исходя из следующего выражения:

$$C_3 = K_{yn}(C_o) + K_3(C_o) + K_p(C_o) + C_{y\partial}R(P - t_a) + R_o(T_a) \times \\ \times (17,5C_m + C_m) + \sum_{l=1}^L N_l C_l + K_c(C_{cm}) \quad , \quad (10)$$

где K_{yn} , K_3 , K_p , K_c – нормативные показатели годовых расходов на услуги промышленности, восполнение агрегатов и ЗИП, средние и капитальные ремонты и эксплуатацию объектов капитального строительства соответственно; $C_{y\partial}$ – стоимость одного киловатт-часа электроэнергии; R – максимальная мощность, потребляемая образцом вооружения; t_a – среднее время работы образца от автономных источников питания в течение года; R_o – мощность первичных источников питания; C_m , C_m – стоимость одного килограмма топлива и смазочных материалов соответственно; N_l – количество специалистов l -ой категории; C_l – годовые расходы на поддержание одного специалиста l -ой категории.

Выводы

Таким образом, методика оценки эффективности системы ТОиР позволяет функционально увязать параметры системы с требуемым составом РТУ, нормами поставки РТУ и затратами на их эксплуатацию.

Методика может использоваться для решения следующих задач:

- оценка эффективности системы ТОиР;
- обоснования требуемых параметров системы ТОиР;
- обоснования направлений совершенствования системы ТОиР.

Библиографический список

1. Назмутдинов И.С. Выбор показателя эффективности системы технического обслуживания и ремонта радиотехнических средств. // Материалы международной научно-практической конференции. – Воронеж: Воронежский институт МВД России. 2014. – С. 80-84
2. Олейников Л.Ф. Эксплуатация и ремонт вооружения и военной техники радиотехнических войск на этапе перевооружения. – М.: Воениздат, 1990. – 216 с.

I.S. Nazmutdinov, S.I. Burobin

ESTIMATION OF THE EFFICIENCY OF THE SYSTEM OF MAINTENANCE AND REPAIR OF RADIO DEVICES

Military Educational and Scientific Center of the Air Force «N.E. Zhukovsky and Y.A. Gagarin Air Force Academy», Voronezh

The article discusses an approach to assessing the effectiveness of the maintenance and repair system, which makes it possible to functionally link the system parameters with the required composition of radio technical devices, the norms for their delivery and the costs of their operation. The system of differential equations of Kolmogorov was chosen as a mathematical apparatus.

Keywords: radio engineering device, maintenance and repair system, Kolmogorov system of differential equations, operation, repair, cost.

ВЫБОР СПОСОБОВ ФОРМАЛИЗАЦИИ ФУНКЦИЙ УПРАВЛЕНИЯ, ПОДЛЕЖАЩИХ АВТОМАТИЗАЦИИ, ПРИ ПРИНЯТИИ РЕШЕНИЙ ДОЛЖНОСТНЫМИ ЛИЦАМИ ОРГАНОВ УПРАВЛЕНИЯ СИСТЕМЫ СВЯЗИ

*Военный учебно-научный центр Военно-воздушных сил
«Военно-воздушная академия имени профессора Н.Е. Жуковского и Ю.А. Гагарина»,
г. Воронеж*

В статье рассмотрен вариант выбора способов формализации функций управления, подлежащих автоматизации, при принятии решений должностными лицами органов управления системы связи, состоящий из алгоритма определения функций управления подлежащих автоматизации и алгоритма обоснования допустимых способов формализации задач управления.

Ключевые слова: автоматизированная система управления, система связи, функции управления, информационные задачи, расчетные задачи, математические модели, экспертная оценка.

Введение

Опыт внедрения и работы автоматизированных систем управления (АСУ) различного назначения показывает, что при автоматизации управления в основном сохраняются последовательность и содержание работ управляющих органов, но изменяется методика и технология управления. АСУ позволяет автоматизировать основные трудоемкие процессы и освободить личный состав для разносторонней оценки вариантов и принятия оптимальных решений по заданным критериям. При этом отпадает функция книжного учета наличия, состояния и движения средств, сводятся к минимуму трудоемкие и ручные расчеты, появляются функции выбора из общего перечня задач, конкретно необходимых в данных условиях, подготовки информации для решения на ЭВМ, определения порядка выдачи выходной информации и ее передачи исполнителям, а также организация размещения и обслуживания средств автоматизации.

Главной целью, которая ставится при автоматизации управления системы связи, является повышение эффективности управления за счет обеспечения оперативности, полноты и достоверности получения информации, оптимальности методов управления и оперативности доведения управленческих решений до исполнителей.

Автоматизация управления в системе связи требует наличия определенного специального программного обеспечения (СПО). Под СПО понимается совокупность программных средств, осуществляющих смысловую обработку информации в интересах деятельности соответствующих органов управления. СПО является сугубо специфичным для каждого органа управления. В его состав должны входить комплексы расчетных задач, автоматизирующие определенные функции, выполняемые органом управления. Поэтому при разработке СПО важное значение имеют работы по формированию перечня расчетных задач и взаимная их увязка в соответствующие комплексы, предназначенные для обработки информации в ходе автоматизированного выполнения соответствующей функции управления.

Основная часть

Исходными данными для разработки способов формализации функций управления, подлежащих автоматизации, при принятии решений должностными лицами органов управления системы связи являются:

- структура органов управления
- множество функций управления, подлежащих автоматизации;

- множество задач управления, подлежащих автоматизации;
- объем информации, предназначенный для решения задач;
- способы формализации задач.

Алгоритм определения функций управления $\{\Phi_j\}$, $j = \overline{1, J}$ подлежащих автоматизации и формирование перечня задач $\{Z_{ji}\}$, $j = \overline{1, J}$, $i = \overline{1, I}$ по каждой функции управления представлен на рисунке 1.



Рис.1. Алгоритм определения перечня задач управления, подлежащих автоматизации

Проведя анализ функционирования органов управления при автоматизации, могут быть автоматизированы следующие функции управления [1]:

1. Функции, связанные со сбором, обработкой, учётом и выдачей информации, необходимой для управления составом, ресурсом и состоянием.
2. Функции, связанные с разработкой документов по планированию решения задач управления.
3. Функции, связанные с доведением принятых решений до исполнителей.

После определение функций управления, подлежащих автоматизации и формирования перечня задач по каждой функции управления переходим к обоснованию допустимых способов формализации задач управления.

По способам формализации задачи могут быть условно разделены на три класса: информационные задачи, расчетные задачи и математические модели.

Общим для класса **информационных задач** является наличие большого объема перерабатываемой информации и логический характер процесса решения. Задачи характеризуются тем, что в ходе решения фиксируется или выдается информация, как правило, в том виде, в каком она вводилась в систему. При этом применяются простейшие способы обработки (обобщение данных, преобразование общих показателей в другие и т.д.).

Класс **расчетных задач** характеризуется относительно небольшим объемом перерабатываемой информации, но сложными и разнообразными методами решения. Отличительной особенностью задач этого класса является то, что в процессе решения получается новая информация, которая в явном виде в систему не вводилась.

Математические модели обеспечивают анализ и оценку информационной ситуации, выработку рекомендаций для принятия решения. В ходе преобразования осуществляется воспроизведение поведения системы во времени и (или) пространстве.

Алгоритм обоснования допустимых способов формализации задач управления представлен на рисунке 2.

Выбор допустимых способов формализации задач, предназначенных для автоматизации процессов управления, осуществляется методом непосредственной экспертной оценки [2]. На начальном этапе осуществляется формирование группы экспертов.

Исходной информацией для выбора способов формализации является:

множество функций управления, подлежащих автоматизации;

множество задач управления, подлежащих автоматизации;

способы формализации;

объем информации, предназначенный для решения задач управления.

В начале осуществляется процедура соответствия имеющегося объема информации k -ой задачи управления при v -ом способе формализации задачи, элементы которого определяются следующим образом:

$$w_{kv} = \begin{cases} 1, & \text{если } k\text{-ая задача может быть формализована } v\text{-ым способом} \\ 0, & \text{иначе} \end{cases}$$

$$v = \begin{cases} 1, & \text{информационная задача} \\ 2, & \text{расчетная задача} \\ 3, & \text{математическая модель} \end{cases},$$

$$\text{при этом } \sum_{v=1}^3 w_{kv} = 1$$

Полученная матрица оформляется в виде таблицы – соответствия имеющегося объема информации k -ой задачи управления при v -ом способе формализации, в которой для каждой задачи управления единичное значение в графах соответствует имеющийся объем при v -ом способе формализации.

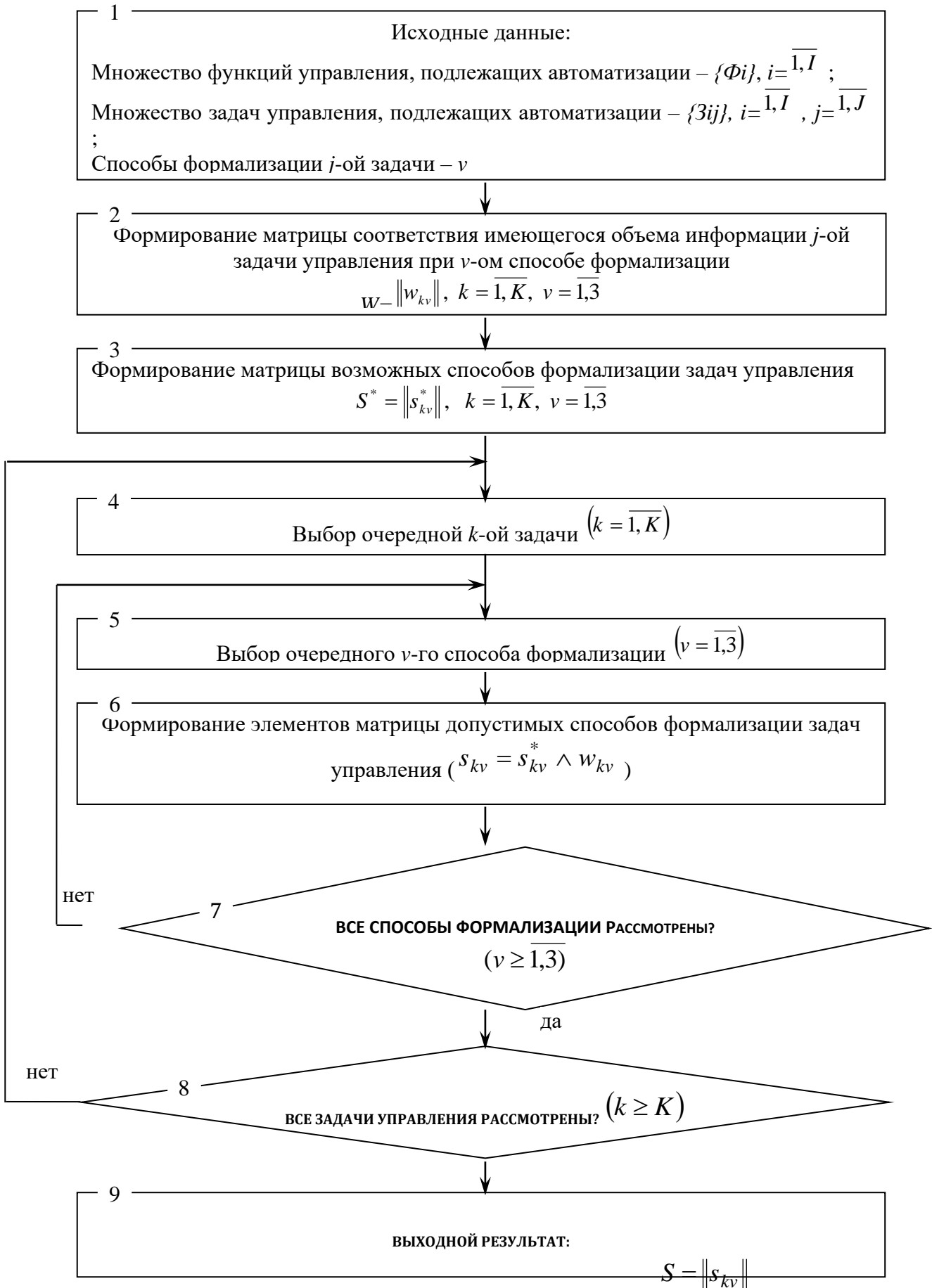


Рис. 2. Алгоритм обоснования допустимых способов формализации задач управления

На следующем шаге с помощью метода непосредственной экспертной оценки, определяется матрица возможных способов формализации задач управления, с учетом целесообразности того или иного способа.

По результатам обработки для каждой из задач должны быть определены возможные способы ее формализации из трех возможных:

- формализация в виде информационной задачи;
- формализация в виде расчетной задачи;
- формализация в виде математической модели.

В процессе оценки каждый эксперт определяет возможные способы формализации задач управления из имеющегося перечня. При этом если возможно формализовать в виде информационной, расчетной или математической модели, то каждому возможному способу формализации присваивается условная оценка I . Если задачу нельзя формализовать тем или иным способом, то такому способу присваивается оценка 0 .

По результатам опроса каждого эксперта формируется матрица оценок возможных способов формализации задач ($S^* = \lceil \lceil s_{kv}^* \rceil \rceil$) без учета имеющихся для их решения объемов, которая оформляется в виде таблицы.

Средняя оценка возможного способа формализации для каждой задачи рассчитывается по формуле:

$$S_{kv}^* = \left\lceil \sum_{l=1}^L x_{kvl} \times k_l \right\rceil, \quad k = \overline{1, K}, \quad v = \overline{1, 3}, \quad (1)$$

где x_{kvl} – оценка l -го эксперта по k -ой задаче; k_l – коэффициент компетентности l -го эксперта; L – количество экспертов; $\lceil \rceil$ – операция округления до ближайшего целого.

Далее формируются элементы матрицы допустимых способов формализации задач:

$$S_{kv} = s_{kv}^* \wedge w_{kv}. \quad (2)$$

Выводы

Таким образом, полученная матрица оформляется в виде таблице, и может использоваться для обоснования требований к основным тактико-техническим характеристикам автоматизированной системе управления.

Библиографический список

1. Военная кибернетика: методология создания автоматизированных систем управления техническим обеспечением. / Под ред. Ю.И. Арпина. – Тверь: ЗАО НИИ ЦПС. – 2006. – 204 с.
2. Сиротинин Е.С., Малышев А.Н. Экспертные оценки в теории принятия решений. – Тверь: ВА ПВО, 1991. – 45 с.

I.S. Nazmutdinov, E.P. Kaydash

**SELECTION OF METHODS FOR FORMALIZING CONTROL FUNCTIONS
TO BE AUTOMATED WHEN MAKING DECISIONS BY OFFICERS
OF COMMUNICATION SYSTEM CONTROL BODIES**

Military Educational and Scientific Center of the Air Force «N.E. Zhukovsky and Y.A. Gagarin Air Force Academy», Voronezh

The article considers the option of choosing ways to formalize control functions to be automated, when making decisions by officials of the control bodies of a communication system, consisting of an algorithm for determining control functions to be automated and an algorithm for justifying acceptable ways to formalize control tasks.

Keywords: automated control system, communication system, control functions, information tasks, computational problems, mathematical models, expert assessment.

СПОСОБ АДАПТИВНОГО УПРАВЛЕНИЯ ВРЕМЕННЫМИ ПАРАМЕТРАМИ АВТОМАТИЗИРОВАННОЙ СЕТИ ДКМ РАДИОСВЯЗИ

*Военный учебно-научный центр Военно-воздушных сил «Военно-воздушная академия
имени профессора Н.Е. Жуковского и Ю.А. Гагарина», г. Воронеж*

Предложен способ повышения разведзащищенности сети радиосвязи на основе минимизации времени излучения передатчиков радиостанций, структурно объединенных в автоматизированную радиосеть для параллельной ретрансляции сигналов. Оптимизация параметров элементов сети, при использовании способа адаптивного управления временными параметрами, обеспечивает повышение ее разведывательной защищенности.

Ключевые слова: разведзащищенность, декаметровая радиосвязь, адаптивное управление, параллельная ретрансляция, попеременное излучение.

Введение

Основной задачей системы связи авиационного формирования является своевременная передача сообщений, с требуемой достоверностью и безопасностью. К системе связи предъявляются требования к наиболее существенными, с точки зрения пользователя, свойствам, которыми являются боевая готовность, устойчивость, пропускная способность, разведзащищенность, мобильность, управляемость и доступность [1].

В ходе боевых действий, пункты управления авиационных формирований подвергаются интенсивному воздействию комплексов радиоэлектронной борьбы (РЭБ) противника на этапе оперативного развертывания, на этапе ведения боевых действий, а также на этапе восстановления боеспособности.

Вскрытые средствами радио и радиотехнической разведки (РПТР) средства радиосвязи, являющиеся основными средствами управления авиацией, позволяют противнику вскрыть систему связи и пункты управления авиационного формирования и принять решение на их огневое или радиоэлектронное подавление.

Таким образом, в условиях воздействия комплексов РЭБ противника, актуальной задачей является обеспечение требуемой разведзащищенности системы радиосвязи.

Выбор показателей разведзащищенности системы связи

В качестве показателя разведзащищенности системы связи может быть выбран коэффициент разведзащищенности $K_{рз}$, определяющий способность системы связи противостоять мерам РПТР, направленным на обнаружение факта работы элементов системы связи, определения параметров сигнала необходимых для радиопротиводействия и вскрытия структуры системы управления [2].

Задача разведки параметров элементов системы связи, как правило, разбивается на несколько последовательно решаемых задач: обнаружения радиоизлучений, местоопределение источника радиоизлучения (ИРИ), радионаблюдение, определение типа ИРИ и их экземплярного распознавания. Причем, если излучающее средство не обнаружено, то оценка его параметров теряет всякий физический смысл. В соответствии с указанным вероятность радиоразведки РЭС целесообразно представить произведением [3]:

$$P_{рр} = P_{обн} \cdot P_{к} \cdot P_{а} \cdot P_{пел} \cdot P_{оп}, \quad (1)$$

где $P_{обн}$ – безусловная вероятность энергетического обнаружения ИРИ; $P_{к}$ – условная вероятность установления контакта постом поиска с ИРИ; $P_{а}$ – условная вероятность анализа характеристик РЭС постом поиска; $P_{пел}$ – условная вероятность

правильного измерения пеленга на ИРИ; $P_{оп}$ – условная вероятность опознавания источника радиосигнала.

Указанные вероятности являются частными показателями разведзащищенности системы связи и, в общем случае, зависят как от энергетических, так и временных характеристик сигналов. При этом энергетические показатели должны характеризоваться энергетическими соотношениями мощностей или уровней сигналов и помех на входе радиоприемного устройства (РПУ), а временные показатели – отражать динамику поведения во времени средств связи и систем РРТР.

Таким образом, показатели разведзащищенности системы связи, для дальнейшего анализа способов повышения разведзащищенности, удобно разделить на показатели энергетического и вероятностно-временного характера.

В соответствии с этим коэффициент разведзащищенности $K_{рз}$ представим в виде [3]:

$$K_{рз} = 1 - P_{обн} \cdot P_k \cdot P_a \cdot P_{пел} \cdot P_{оп} , \quad (3)$$

Задача повышения разведзащищенности системы связи связана с минимизацией частных показателей $P_{обн}$, P_k , P_a , $P_{пел}$ и $P_{оп}$, входящих в выражение (3). Решение данной задачи на основе комплексной оптимизации энергетических ($P_{обн}$, $P_{пел}$), временных (P_k , P_a), структурных и признаковых показателей ($P_{пел}$) элементов системы связи позволит обеспечить требуемое (максимальное) значение коэффициента разведзащищенности $K_{рз}$ системы связи в целом [3].

Критерий оценивания качества системы связи по показателю разведзащищенности может быть представлен в виде:

$$K_{рз} \geq K_{рз доп} , \quad (4)$$

где допустимое значение $K_{рз доп}$ определяется требованиями руководящих документов, либо на основании метода экспертных оценок.

От показателя разведзащищенности зависят другие показатели свойств системы связи: устойчивость (живучесть и помехозащищенность); управляемость; доступность и боевая готовность, следовательно, данный показатель во многом определяет качество системы связи.

С другой стороны, обеспечение допустимого значения коэффициента разведзащищенности не должно приводить к срыву основной задачи системы связи из-за ухудшения своевременности, достоверности и безопасности связи.

Поскольку показатель разведзащищенности, как следует из выражения (3), зависит от энергетических и временных показателей, при разработке способов повышения качества связи по разведзащищенности, целесообразно обеспечить требуемые значения достоверности и своевременности связи. При этом, безопасность связи, как правило, обеспечивается применением аппаратуры криптографической защиты во вторичных сетях и мероприятия обеспечения разведзащищенности значительного влияния на безопасность связи не окажут.

Таким образом, для решения задачи повышения разведзащищенности требуется решение частных задач:

- обеспечение требуемой достоверности связи;
- минимизация отношения сигнал/помеха на входе приемника системы РРТР;
- минимизация времени работы радиолинии на одной частоте;
- минимизация информативности демаскирующих признаков (ДМП).

В работе [4] предложен метод синфазного сложения мощности передатчиков, использующихся для параллельной ретрансляции сигналов пространственно распределенными элементами автоматизированной сети радиосвязи.

Данный метод позволил разработать способ повышения разведзащищенности на основе оптимизации энергетических показателей ($P_{обн}$, $P_{пел}$) при минимизации мощности объединяемых для параллельной ретрансляции передатчиков пропорционально их количеству. Это, в свою очередь позволяет повысить разведзащищенность сети радиосвязи при обеспечении требуемой достоверности [5].

Дальнейшее повышение разведзащищенности предлагается на основе разработки способа оптимизации временных показателей (P_k , P_a), входящих в выражение (3), при минимизации времени излучения передатчиков, структурно объединяемых для ретрансляции сигналов.

Оптимизация временных показателей разведзащищенности автоматизированной сети ДКМ радиосвязи на основе минимизации времени излучения передатчиков

Рассмотрим автоматизированную сеть ДКМ радиосвязи, состоящую из узла отправителя сообщения, узла получателя сообщения и четырех узлов-ретрансляторов (радиостанций). Вариант подобной системы представлен на рисунке 1.

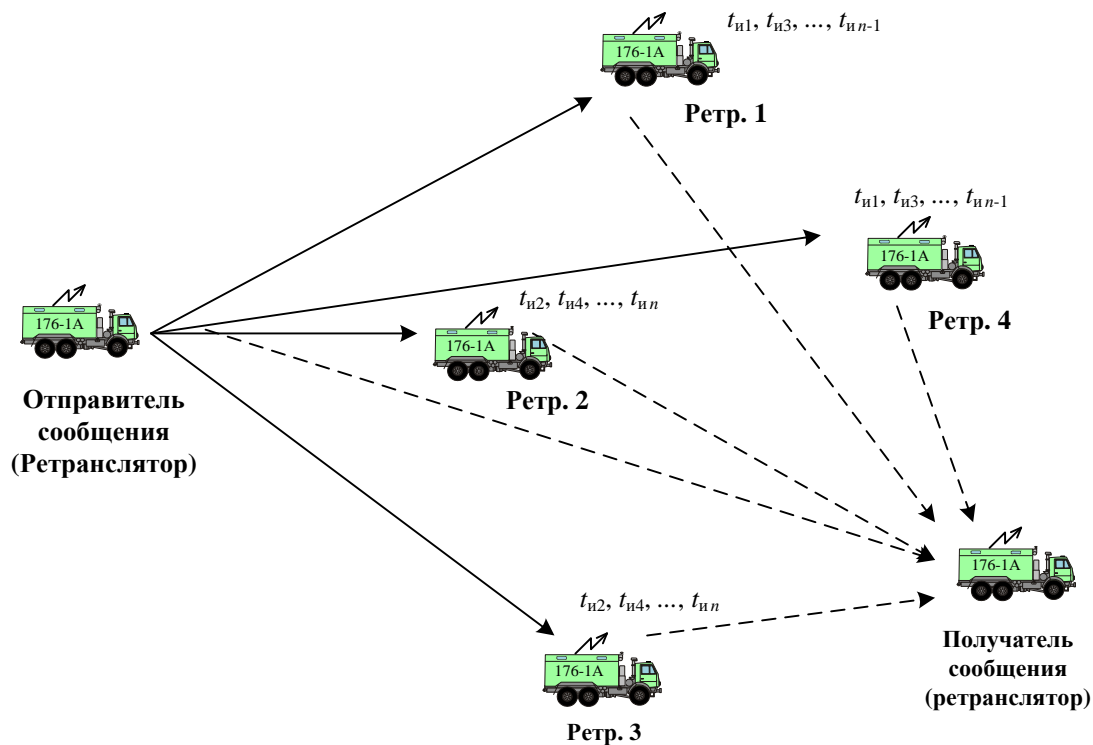


Рис. 1. Вариант автоматизированной сети ДКМ радиосвязи с возможностью попеременного излучения отдельных ИРИ

В общем случае каждый из элементов рассматриваемой радиосети может быть источником (отправителем) сообщений, получателем сообщений или ретранслятором. При этом возможны следующие варианты применения элементов данной радиосети:

- непосредственная передача информации от источника получателю информации;
- последовательная ретрансляция информации от источника получателю через составные радиолинии [6];
- параллельная ретрансляция информации от источника получателю с синфазным сложением сигналов в точке приема [4, 5].

С точки зрения энергетической эффективности наиболее оптимальная структура параллельного ретранслятора должна включать от четырех до шести элементов [4, 5]. Однако, применение даже двух ретрансляторов позволяет более чем в 2 раза уменьшить излучаемую мощность передатчиков и снизить энергетические показатели систем РРТР (вероятность обнаружения $P_{обн}$ и вероятность пеленгации $P_{пел}$ [4, 5].

Рассмотрим возможность снижения временных показателей разведзащищенности и помехозащищенности: вероятности контакта P_k и анализа характеристик P_a на основе минимизации времени излучения элементов автоматизированной сети радиосвязи с адаптивным управлением структурно-параметрическим ресурсом.

Минимизация времени излучения $t_{и}$ предлагается на основе попеременной ретрансляции сигналов информации от источника получателю системой пространственно распределенных ретрансляторов (радиостанций). Для рассматриваемого примера (рисунок 1), наиболее оптимальной является попарная параллельная ретрансляция, обеспечивающая снижение как энергетических ($P_{обн}$, $P_{пел}$), так и временных (P_k , P_a) показателей.

Вариант распределения времени излучения между ретрансляторами для рассматриваемого примера приведен на рисунке 2 а.

Общее время передачи информации $t_{и\Sigma}$ разбивается на интервалы $t_{иi}$, (где $i = 1, 2, \dots, N$) которые распределяются между четырьмя ретрансляторами, осуществляющих попарную попеременную параллельную ретрансляцию (рисунки 1, 2).

Величина интервала $t_{иi}$ может определяться характеристиками системы РРТР противника и, например, удовлетворять требованию:

$$t_{иi} < t_a; \quad (5)$$

где t_a – время анализа характеристик ИРИ после его обнаружения.

При уменьшении доступных ретрансляторов, рассматриваемая сеть радиосвязи должна иметь возможность адаптивно, в автоматизированном режиме перестраивать свою структуру. При этом возможны следующие комбинации попеременно излучающих ретрансляторов:

- при наличии трех ретрансляторов: Ретр.1 и Ретр.2 – Ретр.2 и Ретр.3 – Ретр.3 и Ретр.1 – ... (рисунок 2 б);
- при наличии двух ретрансляторов: Ретр.1 – Ретр.2 – Ретр.1 – ... (рисунок 2 в).

При увеличении количества доступных ретрансляторов, позволяют увеличить количество возможных комбинаций поочередно работающих групп ретрансляторов.

Как следует из анализа результатов моделирования [4, 5], более четырех одновременно излучающих ретранслятора в группе использовать нецелесообразно.

На рисунке 3 представлена зависимость коэффициента разведзащищенности от времени излучения передатчика ДКМ радиостанции, мощностью излучения 1000 Вт.

Анализ результатов расчета, представленной на рисунке 3 зависимости показывает, что при времени излучения менее 5 с коэффициент разведывательной защищенности превышает значение 0,8. При этом, попеременная ретрансляция двух и более ретрансляторов (рисунок 2) позволит пропорционально снизить мощность их передатчиков [5].

Смена групп ретрансляторов может осуществляться на основе одного из трех методов управления: программного управления; управления с обратной связью; адаптивного управления [7].

Программное управление строится в соответствии с жестко заданной программой действий, независимо от реакций объекта управления [7]. Примером программного управления может служить работа радиостанций в режиме ППРЧ. В рассматриваемом способе возможно использовать управление попеременным излучением групп ретрансляторов по псевдослучайному закону, задаваемому программой, которая, по аналогии с ППРЧ, может вводиться в радиостанции в виде ключевых документов.

Недостатком данного метода является нечувствительность системы к изменению радиочастотной обстановки, состоянию радиосети и реакции системы РЭБ противника.

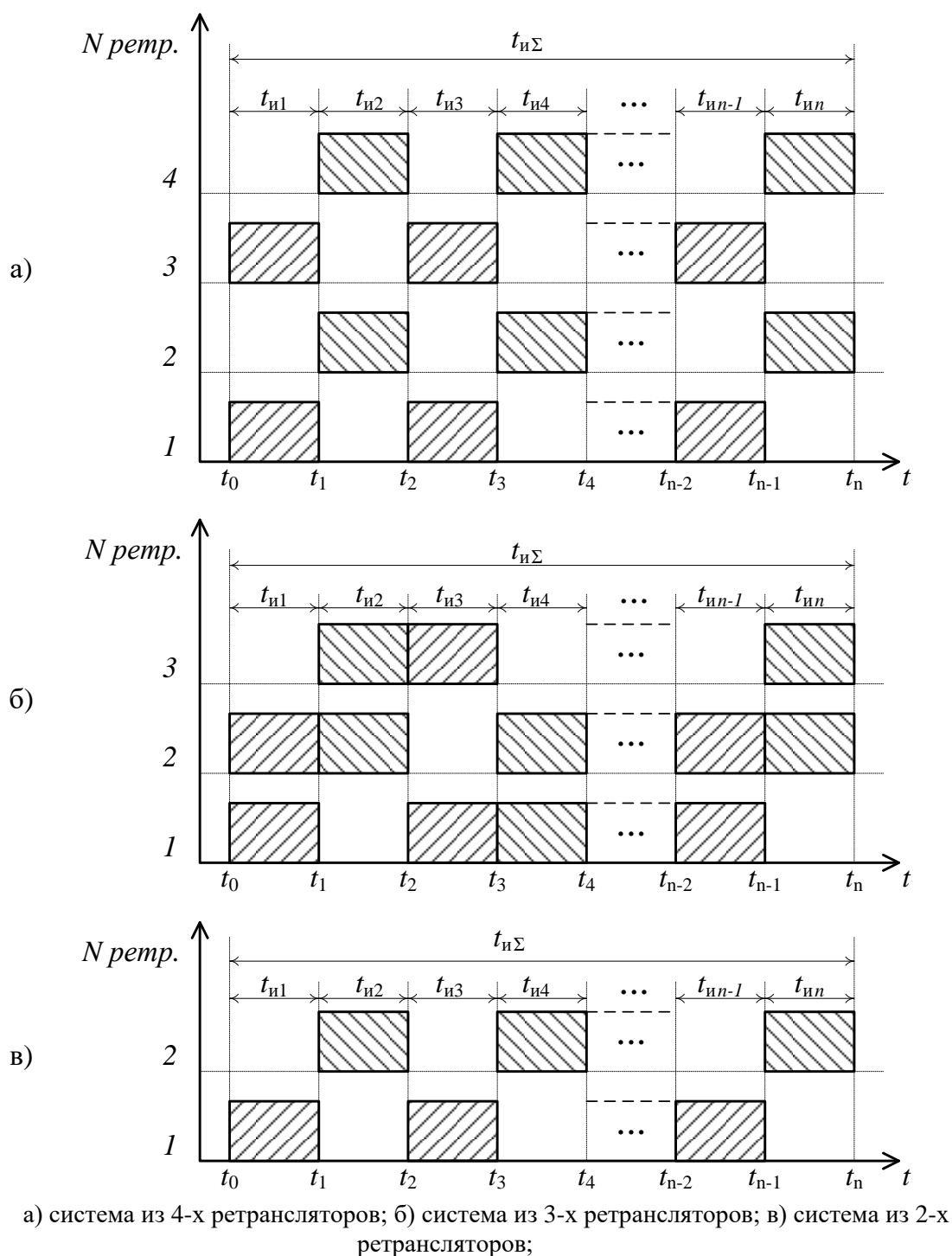


Рис. 2. Вариант распределения времени излучения между ретрансляторами автоматизированной сети радиосвязи с параллельной ретрансляцией сигналов

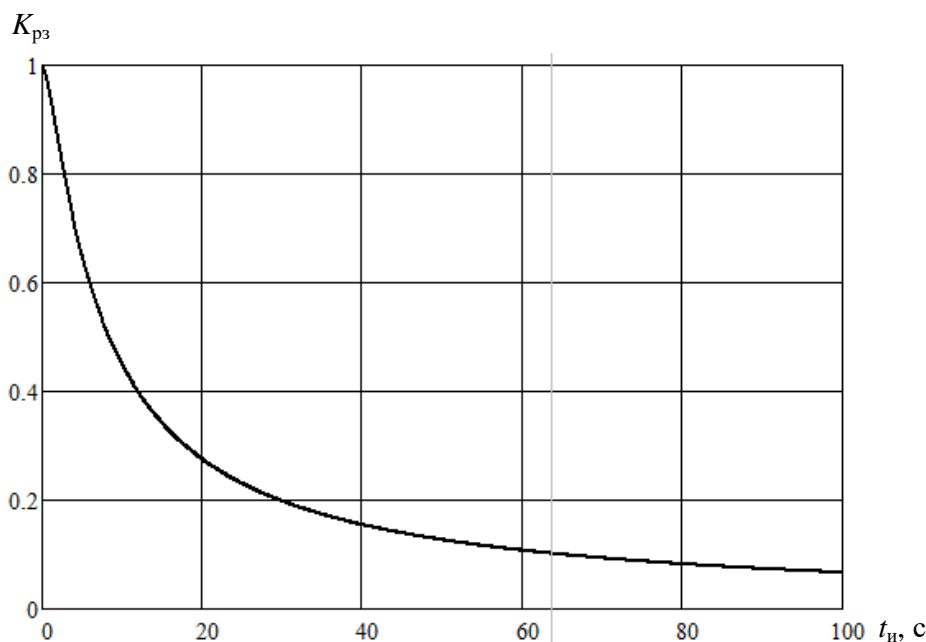


Рис. 3. Зависимость коэффициента разведывательной защищенности от времени излучения

Адаптивное управление строится для «неточно» заданного объекта управления при не полностью известных характеристиках системы РЭБ, и, быть может «нечетко» заданных критериях управления [7]. В рассматриваемом способе структура автоматизированной сети ДКМ радиосвязи и ее режимы работы зависят от радиочастотной обстановки и реакции противоборствующей системы РЭБ. Следовательно, адаптивное управление характеризуется наилучшим качеством.

Выводы

Разработанный способ адаптивного управления временными параметрами автоматизированной сети ДКМ радиосвязи, позволяет минимизировать время излучения радиоприема на одной частоте на основе попеременной ретрансляции сигналов пространственно распределенными радиостанциями, объединяемых в группы для параллельной ретрансляции сигналов с синфазным сложением мощности излучения в точке приема. Это, в свою очередь, позволяет оптимизировать временные показатели разведзащищенности системы связи и, в совокупности с оптимизацией энергетических показателей, обеспечить максимальное значение коэффициента разведзащищенности (2).

Библиографический список

1. Ермишян А.Г. Теоретические основы построения систем военной связи в объединениях и соединениях. Ч. 1. Методологические основы построения организационно-технических систем военной связи / А.Г. Ермишян – С. Пб.: ВАС, 2005. – 740 с.
2. Боговик А.В. Эффективность системы военной связи и методы ее оценки / А.В. Боговик, В.В. Игнатов – С.Пб.: ВАС, 2006. – 184 с.
3. Пресняков М.Ю. Повышение разведзащищенности системы авиационной радиосвязи в условиях радиоэлектронного противодействия / М.Ю. Пресняков, А.В. Рябов // Охрана, безопасность, связь. – Воронеж: Воронежский институт МВД России, 2017. – № 1-3. – С. 255 – 259.
4. Рябов А.В. Особенности пространственного сложения мощности передатчиков декаметрового диапазона / А.В. Рябов // Телекоммуникации: Ежемесячный производственный, научно-технический, информационно-аналитический, и учебно-методический журнал. –М.: Наука и технологии, 2012. – № 9. – С. 22 – 28.

5. Рябов, А.В. Способ повышения разведзащищенности сети декаметровой радиосвязи на основе минимизации мощности передатчиков при параллельной ретрансляции / А.В. Рябов, М.Ю. Пресняков // Сборник материалов Всероссийской военно-научной конференции «Проблемы совершенствования форм и способов применения войск ПВО-ПРО на современном этапе». Секция №7 «Проблемы автоматизации управления войсками ПВО-ПРО на современном этапе». Под общ. ред. А.Н. Тезикова – Тверь: Военная академия ВКО, 2020. – С. 46-51.
6. Головин, О.В. Системы и устройства коротковолновой радиосвязи / О.В. Головин, С.П. Простов – М.: Горячая линия. Телеком, 2006. – 598 с.
7. Семисошенко, М.А. Управление автоматизированными системами декаметровой связи в условиях сложной радиоэлектронной обстановки / М.А. Семисошенко. – С. Пб.: ВАС, 1997. – 364 с.

M. Yu. Presnyakov, A. V. Ryabov

**METHOD FOR ADAPTIVE CONTROL OF TIME PARAMETERS AUTOMATED
DKM RADIO COMMUNICATION NETWORK**

*Military Educational and Scientific Center of the Air Force «N.E. Zhukovsky
and Y.A. Gagarin Air Force Academy», Voronezh*

A method for increasing the intelligence security of a radio communication network based on minimizing the radiation time of transmitters of radio stations, structurally integrated into an automated radio network for parallel signal retransmission. Optimization of the parameters of network elements, when using the method of adaptive control of time parameters, provides an increase in its reconnaissance security.

Keywords: intelligence protection, decameter radio communication, adaptive control, parallel retransmission, alternating radiation

АЛГОРИТМ РАБОТЫ УНИВЕРСАЛЬНОГО АНТЕНННОГО СОГЛАСУЮЩЕГО УСТРОЙСТВА РАСПРЕДЕЛЕННОГО ТИПА

АО «НПП «Полет», г. Нижний Новгород

Современные телекоммуникационные авиационные системы, работающие в ДКМВ диапазоне, должны обеспечивать надежную работу с широким спектром разнотипных антенн, размещаемых на летательных аппаратах. Одним из вариантов повышения надежности связи является согласование радиочастотных фидеров с антеннами различных типов в широкой полосе частот. От решения этой задачи напрямую зависят достоверность передачи информации и дальность связи, масса и габариты бортовых радиостанций и комплексов радиосвязи.

В статье приведена структурная схема и описан алгоритм работы универсального антенного согласующего устройства распределенного типа, предназначенного для согласования разнотипных бортовых антенн ДКМВ диапазона.

Ключевые слова: антенное согласующее устройство, алгоритм.

Универсальное антенное согласующее устройство распределенного типа (УНАСУРТ) [1] состоит из двух коммутируемых на входе и выходе согласующих контуров (СК) «распределенного» типа, узлов «грубой настройки» 1 и 2 (рисунк 1) и устройства управления, которое осуществляет выбор и подключение дискретных элементов согласующих контуров и узлов УНАСУРТ. За счет соответствующих процедур управления происходит настройка элементов УНАСУРТ на любую частоту в пределах заданного диапазона частот. Коммутация элементов и узлов УНАСУРТ осуществляется с помощью высокочастотных вакуумных выключателей и переключателей поляризованного типа, имеющих два устойчивых состояния, за счет чего уменьшается энергопотребление УНАСУРТ.

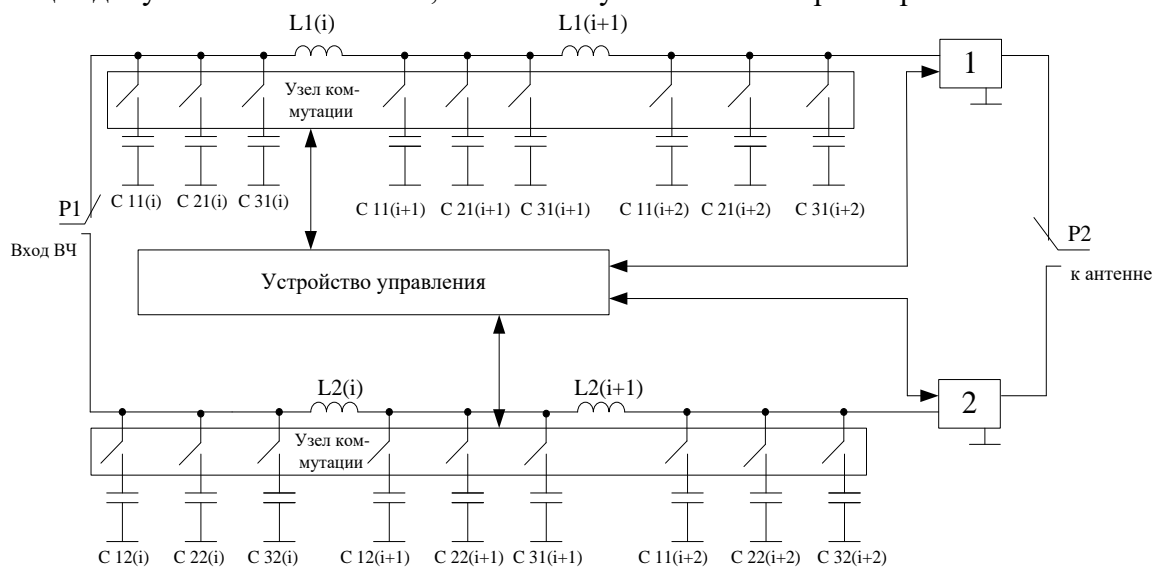


Рис. 1. Структурная схема УНАСУРТ

СК1 и СК2 «распределенного» типа представляют собой отрезки длинной линии, составленные из идентичных П-образных звеньев фильтра нижних частот (ФНЧ), с включенными последовательно катушками индуктивности $L_1(i)$ и $L_2(i)$ и параллельно конденсаторами $C_{11}(i)$ и $C_{12}(i)$, номиналы которых рассчитываются таким образом, чтобы электрическая длина каждого звена была $\theta = \pi/2$.

Предложен линейный алгоритм работы УНАСУРТ, обеспечивающий высокую скорость и точность настройки, достижение максимального значения КБВ на любой частоте в пределах заданного диапазона частот (рисунок 2).

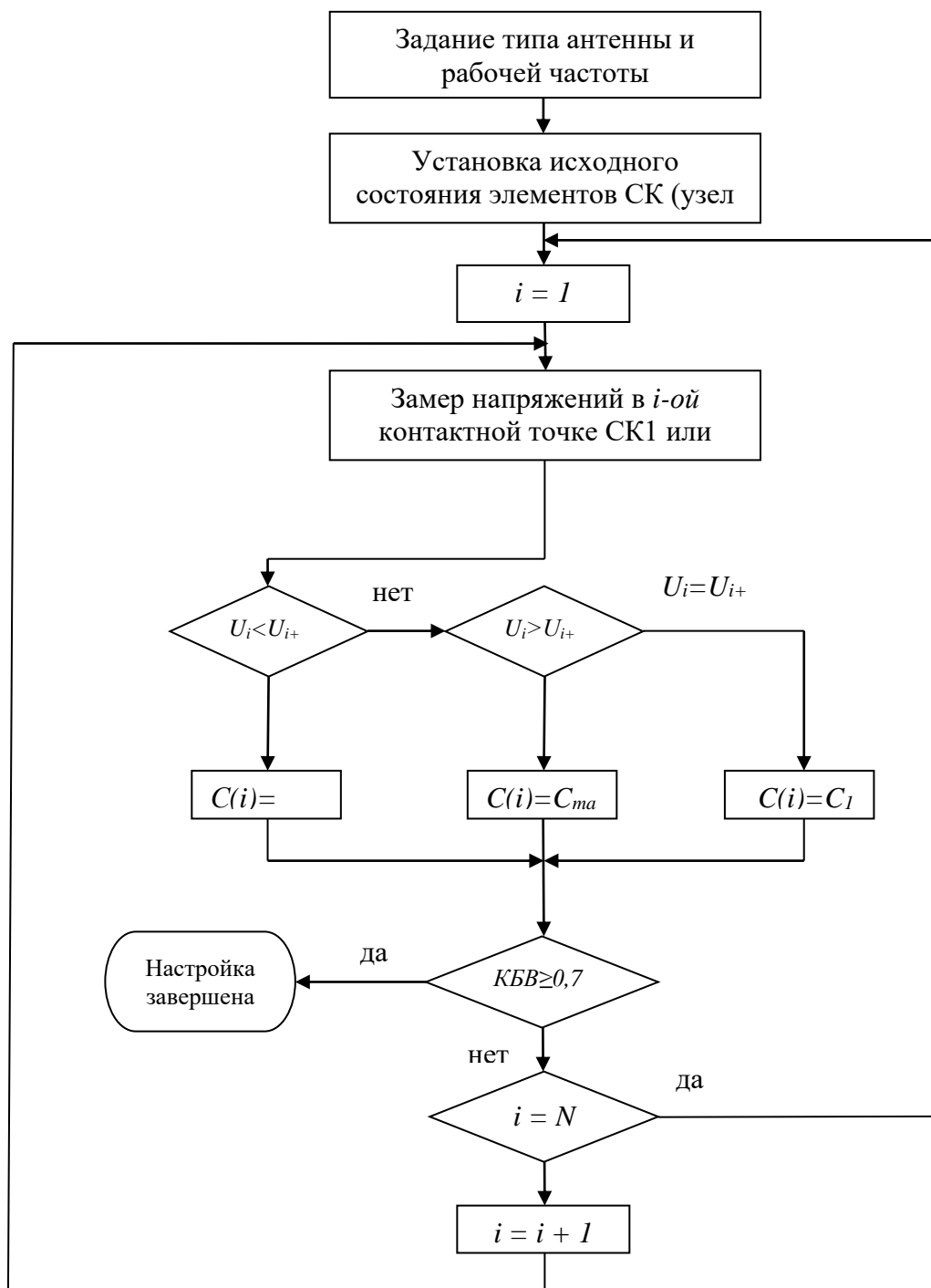


Рис. 2. Блок-схема алгоритма настройки УНАСУРТ

Алгоритм работы УНАСУРТ состоит из двух частей:

-1-я часть: исходя их типа антенны и частоты сигнала коммутируются (при необходимости) элементы узлов 1 или 2;

-2-я часть: работает алгоритм настройки согласующего контура распределенного типа, основанный на пошаговом измерении амплитудных значений высокочастотных напряжений в согласующих контурах, и, в соответствии с полученными значениями, изменением по заданному закону параметров импеданса в контурах следующим образом:

в первом контуре - путем коммутации взамен ранее параллельно подключенных конденсаторов $C_{11}(i)$ конденсаторов $C_{21}(i)=C_{1min}$ или $C_{31}(i)=C_{1max}$; во втором контуре - путем коммутации взамен ранее параллельно подключенных конденсаторов $C_{12}(i)$ конденсаторов $C_{22}(i)=C_{2min}$ или $C_{32}(i)=C_{2max}$.

В процессе работы над диссертацией рассматривались несколько алгоритмов настройки УНАСУРТ, из которых выбран основанный на измерении и сравнении в устройстве управления амплитуд высокочастотных напряжений в соседних ячейках каждого звена подключенного согласующего контура (начиная от антенны), с последующей коммутацией необходимых реактивных элементов согласующего контура следующим образом:

- В случае, если амплитуда напряжения на входе антенны ($U_{ант}$) больше амплитуды напряжения в N -й, ближайшей к антенне, ячейке любого согласующего контура, на величину, большую уровня нечувствительности « δU » сравнивающего устройства, т.е.

$$U_{ант} > U_N + \delta U \quad (1)$$

то посредством устройства управления в N -й ячейке коммутируются емкости $C_{21}(I)=C_{1min}$ или $C_{22}(I)=C_{2min}$.

- В случае, если амплитуда напряжения в i -й ячейке меньше амплитуды напряжения в $(i+1)$ -й ячейке на величину, большую уровня нечувствительности сравнивающего устройства микроконтроллера, т.е.

$$U_i < U_{i+1} - \delta U \quad (2)$$

то посредством устройства управления в i -й ячейке коммутируются емкости $C_{21}(i)=C_{1min}$ или $C_{22}(i)=C_{2min}$.

- В случае, если амплитуда напряжения на входе антенны ($U_{ант}$) меньше амплитуды напряжения в N -й, ближайшей к антенне, ячейке любого согласующего контура, на величину, большую уровня нечувствительности « δ » сравнивающего устройства, т.е.

$$U_{ант} < U_N - \delta U \quad (3)$$

то посредством устройства управления в N -й ячейке коммутируются емкости $C_{21}(I)=C_{1max}$ или $C_{22}(I)=C_{2max}$.

- в случае, если амплитуда напряжения в i -й ячейке больше амплитуды напряжения в $(i+1)$ -й ячейке на величину, большую уровня нечувствительности сравнивающего устройства, т.е.

$$U_i > U_{i+1} + \delta U \quad (4)$$

то посредством устройства управления в i -й ячейке коммутируются емкости $C_{31}(i)=C_{1max}$ или $C_{32}(i)=C_{2max}$.

- в том случае, если напряжения в соседних ячейках равны или отличаются на величину, меньшую уровня нечувствительности сравнивающего устройства, изменения емкости не происходит.

Сравнение напряжений и коммутация необходимых элементов схемы происходит в микроконтроллере устройства управления. Таким образом, происходит настройка на каждую конкретную частоту в пределах рабочего диапазона частот.

Такой алгоритм позволяет быстро (не более 6 с «по алгоритму») и с высокой точностью проводить настройку УНАСУРТ во всем рабочем диапазоне частот. Линейность алгоритма является важнейшим условием стабильности работы УНАСУРТ с антеннами различных типов во всем рабочем диапазоне частот. Алгоритм позволяет работать с простой схемой согласующего контура УНАСУРТ, составленной из одинаковых «распределенных» индуктивностей, набора переключаемых емкостей и устройством управления, выполняющим функцию реализации приведенного выше алгоритма. Данное техническое решение защищено патентом РФ [2].

Устройство управления УНАСУРТ предназначено для автоматического подбора органов настройки элементов СК1, СК2 и узлов 1 и 2, при которых обеспечивается требуемое значение КБВ УНАСУРТ.

Функции УНАСУРТ выполняются либо при взаимодействии аппаратных средств с программным обеспечением (ввод–вывод информации, тест встроенного контроля и другие функции), либо с помощью программного обеспечения – контроль входной информации, формирование выходной информации, выполнение алгоритма управления УНАСУРТ.

Программа настройки УНАСУРТ совместно с аппаратной частью обеспечивает:

- автоматическую настройку УНАСУРТ в рабочем диапазоне частот;
- время настройки УНАСУРТ на произвольной частоте (настройка по алгоритму) – не более 6 с;
- время настройки УНАСУРТ на предварительно настроенных частотах (настройка по памяти) – не более 150 мс;
- работу с радиостанцией в соответствии с протоколом обмена;
- встроенный контроль технического состояния УНАСУРТ.

УНАСУРТ имеет два режима настройки:

- настройка «по алгоритму»;
- настройка «по памяти» на предварительно настроенных частотах.

В режиме настройки УНАСУРТ «по алгоритму» при включении радиостанции или при ее перестройке элементы распределенного СК устанавливаются в исходное состояние (подключены конденсаторы $C_i(i) = C_i$ и катушки индуктивности $L(i)$). Элементы узлов 1 и 2 устанавливаются в зависимости от выбранной рабочей частоты и типа антенны.

Процесс настройки блока завершается, когда из датчика КБВ поступит сигнал, соответствующий достижению на входе УНАСУРТ уровня $КБВ \geq 0,7$. После этого информация о состоянии органов настройки согласующего контура, соответствующая данной частоте, запоминается в репрограммируемом запоминающем устройстве (РПЗУ) устройства управления.

В режиме настройки УНАСУРТ «по памяти», в зависимости от установленной рабочей частоты радиостанции ($F_{раб}$), в устройство управления в последовательном биполярном коде поступает информация о рабочей частоте $F_{раб}$, а также о других частотах: F_n – нижняя граница текущего поддиапазона частот, в пределах которого сохраняется требуемый КБВ, F_v – верхняя граница текущего поддиапазона частот, в пределах которого сохраняется требуемый КБВ, $\Delta F = F_v - F_n$ – размер текущего поддиапазона частот, в пределах которого сохраняется требуемое значение КБВ.

1-й диапазон:	2-й диапазон:	3-й диапазон
$F_n = 2,0 \text{ МГц}$	$F_n = 4,0 \text{ МГц}$	$F_n = 12,0 \text{ МГц}$
$F_v = 3,999 \text{ МГц}$	$F_v = 11,999 \text{ МГц}$	$F_v = 30 \text{ МГц}$
$\Delta F = 1 \text{ кГц}$	$\Delta F = 50 \text{ кГц}$	$\Delta F = 100 \text{ кГц}$

Считанная из РПЗУ информация о включении дискретных элементов настройки согласующего контура, соответствующая установленной рабочей частоте ($F_{раб} - \Delta F \leq F_{раб} \leq F_{раб} + \Delta F$), передается в выходные порты, где она запоминается и поступает далее на входы высокочастотных транзисторов, которые непосредственно управляют высокочастотными реле, коммутирующими элементы согласующего контура.

После включения реле с блока усилителя мощности радиостанции или комплекса связи на вход УНАСУРТ подается настроенный радиосигнал, мощностью (40-50) Вт на время не более 50 мс для оценки уровня КБВ. Если значение $КБВ$ на входе УНАСУРТ оказывается не менее 0,7, процесс настройки заканчивается. В противном случае УНАСУРТ переводится из исходного состояния в режим настройки «по алгоритму» с перезаписью информации в РПЗУ.

В 2015-16 г.г. на предприятии АО «НПП Полет» был изготовлен опытный образец УНАСУРТ мощностью 400 Вт в пике огибающей, и на специализированном технологическом оборудовании проведены настройка и лабораторные испытания на эквиваленты бортовых ДКМВ антенн различных типов.

Библиографический список

1. Хлопушин, И.Ю. К разработке широкополосных адаптивных согласующих устройств для различных типов КВ антенн / И.Ю. Хлопушин // XIX Международная научно-техническая конференция «Радиолокация. Навигация. Связь» - Воронеж – 2013., том 2 – С. 813-819.
2. Патент РФ 2 682 024 С1. Устройство и способ согласования радиопередающего антенно-фидерного тракта с антеннами различных типов / И.Ю. Хлопушин, А.В. Кейстович, Е.С. Анохина – Заявл.10.10.2017. Оpubл. 14.03.2019.

Yu. Khlopushin

ALGORITHM OF OPERATION OF THE UNIVERSAL ANTENNA COUPLER OF THE DISTRIBUTED TYPE

JSC "NPP "Polet", Nizhny Novgorod, RF

Modern telecommunication aviation systems operating in HF band should provide reliable performance with a wide range of different-type antennas installed on aircraft. One of the means to improve communication reliability is matching of radio-frequency feeders with antennas of different types in a wide frequency band. Data transmission validity, communication range, weight and dimensions of on-board radio stations and radio communication complexes directly depend upon the solution of this problem.

This article presents a structure diagram and describes an operation algorithm of all-purpose distributed-type antenna tuning unit designed to provide matching of different-type on-board HF antennas.

Keywords: antenna coupler, algorithm.

ПРОГРАММНЫЙ КОМПЛЕКС «МЕМРИСИМ» ДЛЯ ИМИТАЦИОННОГО МОДЕЛИРОВАНИЯ ИСКУССТВЕННЫХ НЕЙРОННЫХ СЕТЕЙ НА БАЗЕ МЕМРИСТОРОВ

Муромский институт (филиал) Владимирского государственного университета им. А.Г. и Н.Г. Столетовых»

В статье приводится краткое описание авторского программного комплекса «МемриСим» (ПКМС) для имитационного моделирования искусственных нейронных сетей на базе мемристивных устройств (ИНСМ). ПКМС является кроссплатформенным, модульным, расширяемым программным обеспечением. Разработка ПКМС ведется в соответствии с основными принципами системной и программной инженерии, что позволяет организовать работу с ИНСМ на всех этапах их жизненного цикла: проектирования, производства и эксплуатации. В докладе приведен пример работы ПКМС с имитационной моделью ИНСМ, обученной решению задачи распознавания рукописных цифр на датасете MNIST.

Ключевые слова: искусственные нейронные сети, имитационное моделирование, программный комплекс, точность, отказоустойчивость, надежность, производительность.

Введение

Искусственные нейронные сети (ИНС) применяются для решения широкого круга задач, таких как классификация, кластеризация, распознавание образов, аппроксимация и др. [1] В настоящее время одними из наиболее перспективных средств аппаратной реализации ИНС являются мемристивные устройства [2-3]. Они естественным образом подходят на роль искусственных синапсов, потому что могут изменять свое сопротивление в зависимости от прошедшего через них напряжения или тока [4], тем самым изменяя “силу синапса” подобно тому, как это происходит в биологических нейронных сетях.

В процессе создания искусственной нейронной сети на базе мемристивных устройств (ИНСМ) для конкретного проектного решения необходимо определить, какие значения основных показателей качества (точности, отказоустойчивости, надежности, быстродействия, энергопотребления и т.д.) будут в итоге достигнуты. Если принять, что все свойства и характеристики компонентов аппаратно-реализованной ИНСМ будут соответствовать номинальным значениям, то достигнутые значения показателей качества будут зависеть от большого числа параметров (архитектуры, структуры, конструкции, технологии и пр.). На практике, дополнительное влияние на качество будут оказывать неизбежно возникающие дестабилизирующие факторы – шумы и помехи различной природы [5-6]. Это может привести к тому, что созданная ИНСМ не будет соответствовать заданным требованиям, что является значительной проблемой.

На этапе проектирования необходимо определить, какими значениями точности, отказоустойчивости, надежности [7] и других показателей качества будет обладать ИНСМ. Одним из наиболее перспективных средств решения данной задачи является применение методов системной инженерии [8-9] и имитационного моделирования [10-11]. ИНСМ плохо формализуемые или неформализуемые, нелинейные, динамические, многомерные, мультирежимные объекты, с вероятностным характером всех внутренних и внешних явлений и процессов, что в большинстве практических случаев не позволяет применять для их анализа и синтеза только аналитические методы [12].

В настоящее время существует несколько программных продуктов, которые позволяют выполнять компьютерное моделирование ИНСМ, как, например, Memristor

Toolbox [13] для MATLAB и LTSpice [14]. С их помощью можно создавать схемотехнические модели ИНСМ, которые являются необходимыми компонентами проекта и позволяют оценить некоторые показатели качества в номинальном режиме работы, однако с точки зрения системной инженерии этого недостаточно. Для системного проектирования необходимо иметь набор моделей разного уровня структурно-функциональной иерархии, наиболее полно описывающий объект ИНСМ.

Для этих целей авторы разработали программный комплекс [6] «МемриСим» (ПКМС). Компоненты ПКМС позволяют создавать «цифровые двойники» ИНСМ, определять, обеспечивать, оптимизировать их основные технические показатели: точность, отказоустойчивость, надежность, производительность как в номинальных условиях, так и в условиях дестабилизирующих воздействий [15-22].

Программный комплекс «МемриСим»

ПКМС представляет собой кроссплатформенное, модульное, расширяемое программное обеспечение (ПО), в основном написанное на языке программирования Python и имеющее графический пользовательский интерфейс (ГПИ). На рис.1 представлены примеры экранных изображений ПКМС.

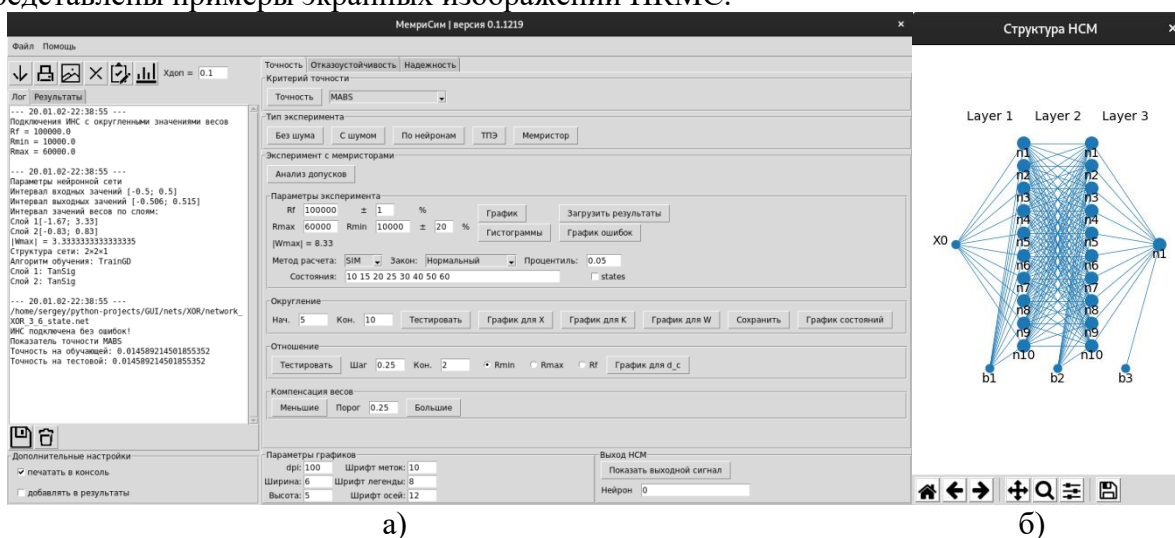


Рис. 1. Экранные изображения программного комплекса «МемриСим»: а) Главное окно программы; б) Пример визуализации структуры многослойного персептрона

Модули ПКМС можно разделить на 2 вида — модули нижнего уровня позволяют выполнять типовые операции над моделями ИНСМ, модули верхнего уровня, используя модули нижнего уровня, позволяют проводить конкретные исследования (например, точности работы, отказоустойчивости или надежности). Для встраивания модуля в ПКМС необходимо сделать для него графический интерфейс модуля (ГИМ), через который можно будет настраивать параметры эксперимента. Модули можно так же использовать без ГИМ, в консольном режиме. На рис. 2 представлена архитектура ПКМС.

Отдельный модуль — «Интерфейс», он необходим для подключения к ПКМС моделей ИНСМ, реализованных в разных программах. Модуль «Интерфейс» содержит класс Connect с определенным набором методов (получить веса из подключённой ИНСМ, установить веса в ИНСМ, симулировать работу ИНСМ в заданный момент времени и т.д.), для которых входные аргументы и выходные параметры определены заранее и используются в других модулях. Поэтому для создания нового «Интерфейса» для работы с моделью, созданной в конкретном ПО, необходимо не меняя имен функций переписать их содержимое. Такой подход обеспечивает универсальность разрабатываемого ПКМС. В настоящее время для него написаны интерфейсы для работы с нейронными сетями, созданными с применением библиотеки NeuroLab и TensorFlow с APIKeras.

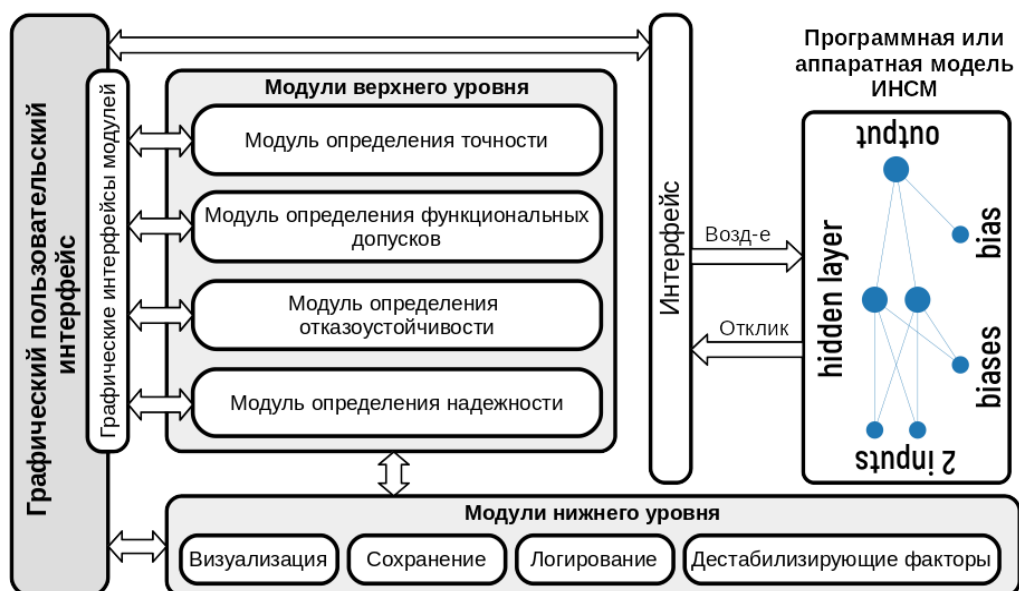


Рис. 2. Архитектура программного комплекса «МемриСим»

Отдельный модуль — «Интерфейс», он необходим для подключения к ПКМС моделей ИНСМ, реализованных в разных программах. Модуль «Интерфейс» содержит класс Connect с определенным набором методов (получить веса из подключённой ИНСМ, установить веса в ИНСМ, симулировать работу ИНСМ в заданный момент времени и т.д.), для которых входные аргументы и выходные параметры определены заранее и используются в других модулях. Поэтому для создания нового «Интерфейса» для работы с моделью, созданной в конкретном ПО, необходимо не меняя имен функций переписать их содержимое. Такой подход обеспечивает универсальность разрабатываемого ПКМС. В настоящее время для него написаны интерфейсы для работы с нейронными сетями, созданными с применением библиотеки NeuroLab и TensorFlow с APIKeras.

Для подключения ИНСМ создан собственный формат файлов «*.training_data». Файлы данного формата в специальных полях должны содержать полную информацию об исследуемой ИНСМ, такую как имя (указывает «Интерфейсу» на подключаемую модель), идентификационный номер, обучающая и тестовая выборки, показатель точности, используемый при обучении и т. д. Для сохранения результатов эксперимента создан собственный формат файлов «*.graph». Файлы данного формата в специальных полях содержат всю информацию о проведенном эксперименте, включая идентификационные данные исследуемой ИНС, raw данные эксперимента и некоторые результаты первичной и (или) вторичной обработки.

Таким образом на данный момент программа обладает следующим функционалом:

- подключение моделей ИНСМ через модуль «Интерфейс» из разных программ, с выводом основных параметров из файла «*.training_data» (интервалы входных и выходных значений, интервал значений весов по слоям, максимальное значение веса, структура сети, алгоритм обучения, пороговые функции по слоям, показатель качества, точность на обучающей и тестовой выборке). При подключении производится тестирование модели и функций системы;
- визуализация структуры ИНСМ, выходного сигнала ИНСМ, гистограммы весов синапсов по слоям, настройка параметров графиков (dpi, ширина, высота, размеры шрифтов)
- задание допуска на показатель качества (точности функционирования);
- определение точности в номинальном режиме по разным критериям (сумма квадратов ошибок, средняя сумма квадратов ошибок, максимальная абсолютная и относительная погрешности, вероятность распознавания, полнота и точность);

- анализ и синтез функциональных допусков (разными методами, при разных законах распределения погрешностей (равномерный, нормальный, Релея и др.));
- моделирование дестабилизирующих факторов;
- планирование эксперимента;
- обработка результатов эксперимента;
- визуализация результатов эксперимента и их аппроксимирование;
- сохранение и загрузка результатов эксперимента.

Описание моделей и эксперимента

Рассмотрим возможность применения ПКМС на примере задачи определения отказоустойчивости ИНСМ, которая создана для распознавания рукописных цифр по их чёрно-белым изображениям размером 28×28 пикселей. Обучающие (50 000 цифр) и тестовые (10 000 цифр) примеры были взяты из базы данных MNIST (рис. 3). Архитектура ИНСМ представляет собой двухслойный персептрон. Количество входов равно количеству пикселей во входном изображении 784. Число выходных нейронов равно числу классов 10. Критерием точности работы ИНСМ является вероятность P распознавания цифры. Допустимая точность для данного варианта исполнения $X_{доп}$ установлена на уровне 90%.



Рис. 3. Примеры рукописных изображений из базы данных MNIST, которая использовалась для обучения и тестирования ИНСМ

Оценка отказоустойчивости будет вычисляться по критерию K .

$$K_i = 1 - \frac{X_i - X_{доc}}{X_{доп} - X_{доc}}, \quad (1)$$

где X_i – значение критерия точности работы ИНСМ, при моделировании погрешности её i -го параметра.

$X_{доc}$ – значение критерия точности ИНСМ, достигнутое при обучении;

$X_{доп}$ – допускаемое значение (допуск) критерия точности работы ИНСМ.

Критерий K является безразмерной величиной и может принимать значения от $-\infty$ до 1. Диапазон изменения K от 0 до 1 характеризует запас точности ИНСМ до достижения границы поля допуска. Если значение показателя X_i , при каких-либо погрешностях параметров ИНСМ становится ниже допускаемого уровня, то она не является работоспособной, то есть произошел отказ, и в этом случае критерий K становится отрицательным.

Эксперимент состоит из следующих этапов:

1. Создание компьютерной модели ИНСМ на функциональном уровне. Сохранение точности работы ИНСМ $X_{доc}$ достигнутой в процессе обучения и оцениваемой на тестовой выборке. Установка допустимого для данной задачи значения точности работы ИНСМ $X_{доп}$ ($X_{доc}$ должен быть больше, чем $X_{доп}$).

2. Проведение эксперимента. Моделирование погрешностей весов синапсов ИНСМ, распределенных по нормальному закону в пределах $\pm \Delta m\%$ с шагом 10%. Нормальное распределение было выбрано в качестве экспериментального. Для реального ИНСМ распределение ошибок синапсов на основе мемристоров может быть установлено эмпирически.

3. Повторение эксперимента $i = 1000$ раз для каждого шага $\pm \Delta m\%$ с сохранением параметров распределения (нижняя граница и верхняя граница) вероятности распознавания изображения P . Вычисление значений критерия отказоустойчивости K (1) для каждого варианта ИНСМ.

4. Интерполяция зависимости между отказоустойчивостью ИНСМ и погрешностью весов синапсов. В рассматриваемом примере использована бикубическая интерполяция. Узлами интерполяции являются экспериментальные значения отказоустойчивости ИНСМ.

5. Определение предельно допустимого значения $\Delta m\%$ ($K = 0$) методами одномерной оптимизации. В данном случае был использован метод золотого сечения.

6. Визуализация результатов эксперимента.

7. Анализ результатов эксперимента и формулировка выводов.

Для проведения исследования были обучены 4 ИНСМ по разным алгоритмам и с разным числом нейронов в скрытом слое: A50-адаптивный алгоритм обучения с оценкой момента, количество нейронов в скрытом слое равно 50; A100-адаптивный алгоритм обучения с оценкой момента, количество нейронов в скрытом слое равно 100; G50 - алгоритм обучения на основе градиентного спуска, количество нейронов в скрытом слое-50; G100 - алгоритм обучения на основе градиентного спуска, количество нейронов в скрытом слое равно 100.

Проведённое исследование точности работы и отказоустойчивости ИНСМ (рис.4, 5) по критерию K показало следующее. Отказоустойчивость ИНСМ зависит как от алгоритма обучения, так и от количества нейронов в скрытом слое.

Результат влияния погрешностей весов синапсов на точность работы ИНСМ показано на рис. 4. Моделирование повторялось 1000 раз для каждого значения Δm с шагом $\pm 10\%$ до тех пор, пока не было достигнуто отрицательное значение критерия отказоустойчивости K . В ходе моделирования были зафиксированы значения вероятности распознавания. Переменная $X_{из}$ формулы (1) критерия отказоустойчивости K равна минимальному значению вероятности распознавания по результатам 1000 экспериментов.

Результат исследования отказоустойчивости ИНСМ показаны на рис. 5. Самую низкую отказоустойчивость имеет сеть G50, а самую высокую -сеть A100. Более точные значения предельно допустимых пределов Δm , полученные методом оптимизации по интерполированным зависимостям, представлены в таблице 1.

Таблица 1.

Результаты исследования точности и отказоустойчивости работы ИНСМ

ИНСМ	После обучения и тестирования			После добавления погрешностей		
	K	$\Delta m, \%$	$P, \%$	K	$\Delta m, \%$	$P, \%$
A50	1.0	0	93.5	0.0	52.2	90.0
A100	1.0	0	93.7	0.0	61.3	90.0
G50	1.0	0	93.5	0.0	48.7	90.0
G100	1.0	0	93.4	0.0	50.3	90.0

Разница в точности работы и отказоустойчивости между ИНСМ со 100 и 50 нейронами в скрытом слое незначительна в соответствии с таблицей 1, но разница в количестве синапсов составляет 2 раза, что в абсолютном значении составляет 39 700 синапсов (или 158800 мемристоров, когда синапс реализован в виде моста из 4

мемристоров). Это показывает, что применение разработанной авторами системы позволяет обосновать выбор архитектуры и структуры, создаваемой ИНСМ не только с точки зрения обеспечения требуемых значений показателей качества, но и с точки зрения определения необходимого и достаточного количества элементов для её физической реализации. Кроме того, обоснованная замена одной ИНСМ на несколько (например, замена A100 на A50 и G50) повышает отказоустойчивость и надежность создаваемой системы и позволяет использовать мажоритарный элемент, соединяющий выходы ИНСМ, тем самым повышая вероятность распознавания.

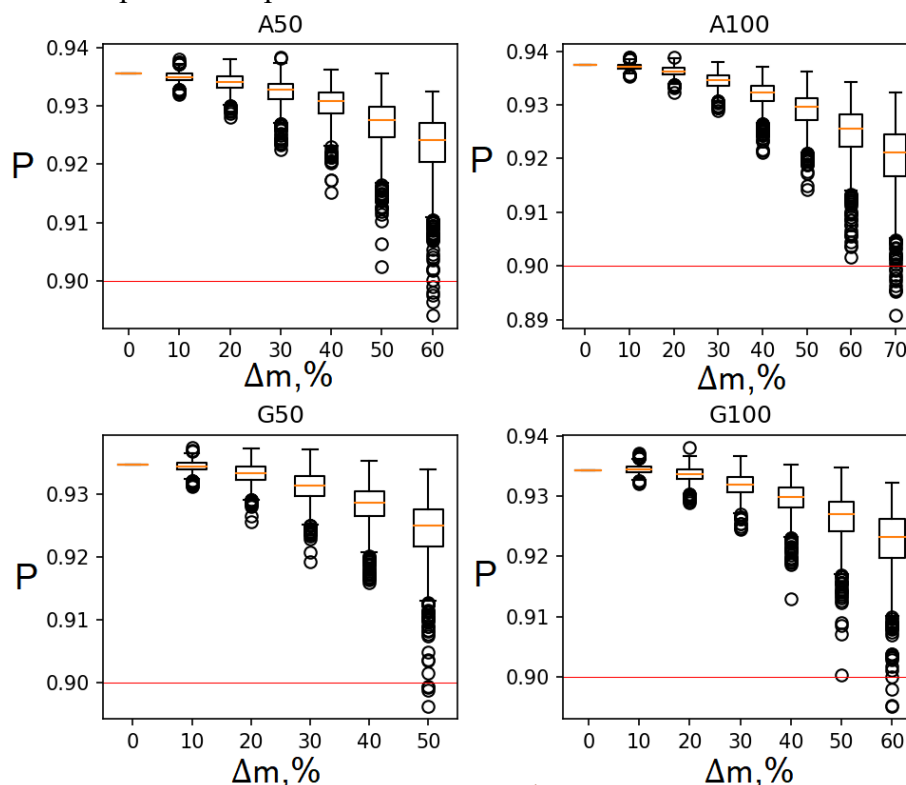


Рис. 4. Исследование точности работы ИНСМ. P -это вероятность распознавания, рассчитанная на тестовом образце. Δm -максимальная относительная погрешность весов синапсов ИНСМ. Красная линия указывает на допустимое значение P

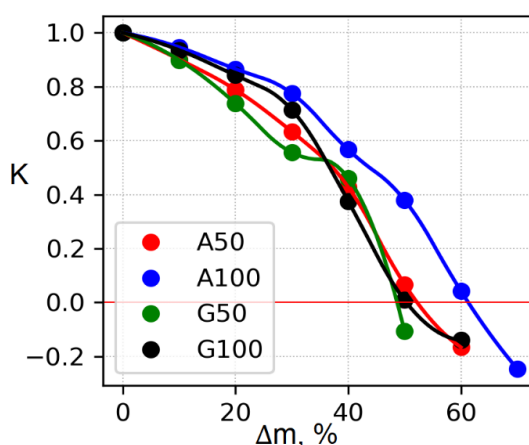


Рис. 5. Исследование отказоустойчивости ИНСМ. Точками обозначены значения отказоустойчивости, полученные при моделировании, линиями - результаты бикубической интерполяции на основе экспериментальных данных

Заключение

Разрабатываемый авторами ПКМС позволяет определять, обеспечивать, оптимизировать основные показатели качества ИНСМ, такие как точность, отказоустойчивость, надежность, производительность как в номинальных условиях, так и в условиях дестабилизирующих воздействий.

Высокая эффективность ПКМС показана на примере оптимального проектирования ИНСМ с заданными техническими параметрами для распознавания рукописных цифр из базы данных MNIST.

Авторы продолжают совершенствовать ПКМС и технологии его работы в части комплексного применения методологии системной инженерии на всех этапах жизненного цикла ИНСМ.

Работа выполнена при поддержке гранта РФФИ №19-07-01215 и Стипендии Президента РФ СП-5411.2021.5.

Библиографический список

1. Галушкин А.И. Нейронные сети: основы теории / А.И. Галушкин. – М.: Горячая линия. – Телеком, 2010. – 496 с.
2. Mikhaylov A, Pimashkin A, Pigareva Y, Gerasimova S, Gryaznov E, Shchanikov S, Zuev A, Talanov M., Lavrov I., Demin V., Erokhin V., Lobov S., Mukhina I., Kazantsev V., Wu H and Spagnolo B (2020) Neurohybrid Memristive CMOS-Integrated Systems for Biosensors and Neuroprosthetics. *Front. Neurosci.* 14:358. doi: 10.3389/fnins.2020.00358
3. Erokhin V. Memristive Devices for Neuromorphic Applications: Comparative Analysis. *BioNanoSci.* 10, 834–847 (2020). <https://doi.org/10.1007/s12668-020-00795-1>
4. Demin V Nekhaev D., Surazhevsky I., Nikiryu K., Emelyanov A., Nikolaev S., Rylkov V., Kovalchuk M. (2021). Necessary conditions for STDP-based pattern recognition learning in a memristive spiking neural network. *Neural Networks.* 134. 64-75. 10.1016/j.neunet.2020.11.005.
5. S. Stathopoulos, A. Serb, A. Khiat, M. Ogorzałek and T. Prodromakis, "A Memristive Switching Uncertainty Model," in *IEEE Transactions on Electron Devices*, vol. 66, no. 7, pp. 2946-2953, July 2019, doi: 10.1109/TED.2019.2918102.
6. Filatov D. O., et al. "Noise-induced resistive switching in a memristor based on ZrO₂ (Y)/Ta₂O₅ stack." *Journal of Statistical Mechanics: Theory and Experiment* 2019.12 (2019): 124026.
7. Милов В.Р. Структурно-параметрический синтез нейросетевых систем обработки информации. Диссертация на соискание ученой степени доктора технических наук / Нижегородский государственный технический университет. Нижний Новгород, 2003. – 310 с.
8. Николенко В. Ю. Системный подход к управлению высокотехнологичными проектами / В. Ю. Николенко. — [б. м.]: Издательские решения, 2019. — 252 с.
9. Кулик С.Д. Элементы системного анализа для студентов старших курсов университета. *Естественные и технические науки.* 2018. № 11 (125). С. 373-377.
10. Хранилов В.П. Методология и технологии интерактивного управления ресурсами технических систем при проектировании. Диссертация на соискание ученой степени доктора технических наук / Нижегородский государственный технический университет. Нижний Новгород, 2007. - 355 с.
11. Баранов В. Г. и др. Методы проектирования информационно-управляющих и телекоммуникационных систем.: М: Издательство "Радиотехника" – 2016. 216с.
12. Борданов И.А., Щаников С.А. Перспектива применения имитационного моделирования при проектировании искусственных нейронных сетей на базе мемристоров // *Алгоритмы, методы и системы обработки данных.* 2019. № 2(40). С. 13-19.
13. Memristor Circuit Investigation through a new Tutorial Toolbox, A. Walsh, R. Carley, O. Feely, A. Ascoli, in *Proc. European Conference on Circuit Theory and Design*, Dresden, September 2013.

14. Biolek Z., Biolek D., Biolková V. "Spice Model of Memristor with Nonlinear Dopant Drift", Radio engineer-ing, vol. 18, no. 2, 2009, pp. 210-214.
15. Данилин С.Н., Щаников С.А., Сакулин А.Е., Борданов И.А. Модуль определения отказоустойчивости для системы имитационного моделирования искусственных нейронных сетей на базе мемристоров. Свидетельство о регистрации программы для ЭВМ RUS 2019611745 22.01.2019.
16. Данилин С.Н., Щаников С.А., Зуев А.Д., Борданов И.А. Модуль определения точности функционирования искусственных нейронных сетей на базе мемристоров для системы имитационного моделирования. Свидетельство о регистрации программы для ЭВМ RUS 2019661246 12.08.2019.
17. Данилин С.Н., Щаников С.А., Зуев А.Д., Борданов И.А. Модуль определения функциональных допусков искусственных нейронных сетей на базе мемристоров для системы имитационного моделирования. Свидетельство о регистрации программы для ЭВМ RUS 2019661251 12.08.2019.
18. Данилин С.Н., Щаников С.А., Борданов И.А., Зуев А.Д. Программа автоматизированного определения отказоустойчивости искусственных нейронных сетей на базе мемристоров при возникновении отказов синапсов и нейронов // Свидетельство о государственной регистрации программы для ЭВМ №2020610434. Внесено в Реестр программ для ЭВМ 14.01.2020.
19. Данилин С.Н., Щаников С.А., Борданов И.А. Программа автоматизированного построения схемы надёжности искусственной нейронной сети // Свидетельство о государственной регистрации программы для ЭВМ №2020619399. Внесено в Реестр программ для ЭВМ 17.08.2020.
20. Щаников С.А., Борданов И.А., Данилин С.Н. Программа для выявления неустойчивых к единичным отказам компонентов искусственных нейронных сетей на базе мемристоров // Свидетельство о государственной регистрации программы для ЭВМ №2020664744. Внесено в Реестр программ для ЭВМ 18.11.2020.
21. Щаников С.А., Борданов И.А., Данилин С.Н. Программа для определения оптимального допуска на программирование мемристоров в составе синапсов искусственных нейронных сетей // Свидетельство о государственной регистрации программы для ЭВМ №2020665228. Внесено в Реестр программ для ЭВМ 11.01.2021.
22. Данилин С.Н., Щаников С.А., Борданов И.А., Зуев А.Д. Программа автоматизированного определения надёжности искусственной нейронной сети. Свидетельство о государственной регистрации программы для ЭВМ №2021610038. Внесено в Реестр программ для ЭВМ 24.11.2020.

I.A. Bordanov, S.N. Danilin, S.A. Shchanikov

MEMRISIM SOFTWARE FOR SIMULATION OF ARTIFICIAL NEURAL NETWORKS BASED ON MEMRISTORS

Murom Institute (branch) "Vladimir State University named after Alexander and Nickolay Stoletovs"

The article provides a brief description of the author's software package "MemriSim" (SPMS) for the simulation of artificial neural networks based on memristive devices (ANNM). SPMS is a cross-platform, modular, extensible software. The development of SPMS is carried out in accordance with the basic principles of system and software engineering, which allows you to organize work with ANNM at all stages of their life cycle: design, production and operation. The report provides an example of the work of the SPMS with a simulation model of the ANNM, trained to solve the problem of recognizing handwritten numbers on the MNIST dataset.

Keywords: artificial neural networks, fault tolerance criteria, memristors, neurocomputers, engineering design, accuracy, fault tolerance, reliability.

ПЕРЕСТАНОВОЧНОЕ ДЕКОДИРОВАНИЕ НА ОСНОВЕ ЦИКЛИЧЕСКИХ СВОЙСТВ ПОСЛЕДОВАТЕЛЬНОСТЕЙ СИМВОЛОВ КОДОВОГО ВЕКТОРА

Нижегородский государственный инженерно-экономический университет, г. Княгинино

Для ускорения процесса декодирования принятого кодового вектора из канала связи на основе перестановочного декодирования, ранее было предложено использование когнитивной карты, в памяти которой хранятся все варианты комбинаций символов для исключения проведения матричных вычислений при перестановках нового кодового вектора. Однако в ряде случаев требуется не оправдано большой объём памяти когнитивной карты. Было выяснено, что все приходящие из канала связи символьные последовательности имеют циклический характер, а свойства орбит циклических перестановок позволяют значительно сократить объём памяти когнитивной карты и, соответственно, время декодирования кодовой последовательности.

Ключевые слова: кодовый вектор, перестановочное декодирование, когнитивная карта, орбиты перестановок.

Введение

В современных высокоскоростных сетях управления объектами крайне важно сократить время цикла управления в реальном масштабе времени с целью защиты команд управления от воздействия различного рода деструктивных факторов природного, техногенного и антропогенного характера. Крайне ограниченный фактор времени не позволяет использовать методы помехоустойчивого кодирования с большой избыточностью. Отсюда вытекает потребность использования коротких блоковых помехоустойчивых кодов с максимальным использованием введённой в код избыточности и предельно кратким по времени декодированием команд управления. Рассмотрение путей решения данной задачи показало, что наиболее оптимальным будет использование модифицированного метода перестановочного декодирования (ПД) с использованием принципов когнитивной обработки данных и свойств цикличности орбит перестановок.

Объект исследования

Для реализации ПД необходимы мягкие решения символов (MPC). Широко известны методы формирования подобных решений для двоичного канала связи в условиях воздействия аддитивного шума [5, 6]. Однако для интенсивно развивающихся оптических линий связи подобные методы не разрабатывались. Считалось, что вероятность ошибки на бит в таких структурах имеют низкие значения и поэтому нет необходимости вводить дополнительные решающие правила в систему обработки сигналов. Однако применение сложных видов модуляции требует применения в таких системах средств помехоустойчивого кодирования, эффективность которых существенно возрастает в условиях использования MPC.

Такие значения целесообразно вырабатывать на базе стирающего канала связи с широким интервалом стирания в соответствии с аналитическим выражением вида:

$$\lambda_i(z) = \left\lfloor \frac{\lambda_{\max}}{\mu\sqrt{E_b}} \times z_i \right\rfloor, \text{ при } 0 \leq z_i \leq \mu\sqrt{E_b}, \quad (1)$$

где λ_{\max} – максимальное значение мягкого решения (задаётся конструктором); μ – значение интервала стирания (обычно $0 \leq \mu < 1$); E_b – энергия сигнала на бит; z_i –

реально зафиксированное значение сигнала (как правило, из-за воздействия деструктивных факторов отличается от номинального) [5]. Условный оператор выработки МРС представляется выражением:

$$\lambda_i(z) = \begin{cases} \lfloor (\lambda_{\max} / \mu \sqrt{E_b}) \times z_i \rfloor, & \text{при } 0 \leq z_i < \mu \sqrt{E_b}; \\ \lambda_{\max}, & \text{при } \mu \sqrt{E_b} \leq z_i < (2 - \mu) \sqrt{E_b}; \\ -\lfloor (\lambda_{\max} / \mu \sqrt{E_b}) \times z_i \rfloor, & \text{при } (2 - \mu) \sqrt{E_b} \leq z_i < 3\sigma. \end{cases} \quad (2)$$

Таким образом, выражение (1) характеризует набор линейных функций, имеющих максимальное значение именно в области математического ожидания обрабатываемых сигналов.

В отличие от известного метода отношений правдоподобий [1] представленное выражение (1) не имеет связи с дисперсией помехи, что позволяет использовать его в оптических системах связи, например, при использовании модуляции РАМ-4. Действительно, в оптических системах при измерении уровня помех вместо дисперсии σ^2 гауссовского шума используется понятие «коэффициента ошибок» и выражение вида $\lambda_i(z) = (2z_i \sqrt{E_b}) / \sigma^2$ для оптических каналов связи теряет смысл.

В новых условиях кодовый вектор длины n в процессоре приёмника фиксируется в виде некоторой последовательности $+\lambda_1, +\lambda_2, -\lambda_3, \dots, +\lambda_k, \dots, -\lambda_{n-1}, +\lambda_n$, в которой знак $+$ соответствует биту со значением 1, а знак $-$ соответствует информационному биту со значением 0. Для реализации процедуры упорядочения символов по значениям их МРС знаки жестких решений значения не имеют. Сортировка завершается, если, например, установлено, что $|\lambda_3| \geq |\lambda_n| \geq |\lambda_5| \geq \dots \geq |\lambda_1| \geq |\lambda_2|$, и на этой основе формируется биекция вида:

$$b = \begin{pmatrix} 1 & 2 & 3 & \dots & n-1 & n \\ \lambda_3 & \lambda_n & \lambda_5 & \dots & \lambda_1 & \lambda_2 \end{pmatrix}. \quad (3)$$

В последнем выражении верхняя строка определяет область определения перестановки, которая в прикладном значении для теории избыточного кодирования совпадает с длиной n кодового вектора выбранного кода C_j , а нижняя строка определяет область значений МРС. В биекции b для элементов нижней строки важно выделить две группы значений МРС, которые будут определять два непересекающихся класса эквивалентности элементов из области определения. В общем виде согласно выражению (4) это представляется в виде произведения двух непересекающихся циклов нумераторов вида $(3 \ n \ 5 \dots) (\dots 1 \ 2)$. Следовательно, группа нумераторов левого цикла представляет множество наиболее надёжных элементов $\{F_{rel}\}$ принятого кодового вектора. Число элементов левого класса эквивалентности определяется значением числа информационных символов k кода C_j , правый класс эквивалентности $\{\overline{F_{rel}}\}$ определяется числом избыточных элементов кода $r = n - k$. Применительно к для каждого принятого вектора формируется две орбиты: орбита X_k – с элементами информационных разрядов ЭК и орбита X_r – с элементами избыточных разрядов ЭК.

Циклические свойства перестановок и их орбиты. Под перестановками элементов (чаще всего чисел) принято понимать все возможные способы, которыми их можно представить. Известно, что если во множестве чисел содержится n элементов, то всего может быть сформировано $n!$ различных перестановок. В системе ПД, как правило, используются блочные систематические (n, k) – коды, где n – общая длина кодовой комбинации, а k – число информационных разрядов в ней, при этом $k < n$. Указанный способ декодирования данных обеспечивает по сравнению с другими известными

алгоритмами обработки избыточных кодов лучшие показатели эффективности по энергетическому выигрышу кода (ЭВК) [1 – 3]. При организации процедуры ПД декодер по какому-либо критерию должен выбрать k надёжных символов из n принятых и в общем случае может быть создано C_k^n таких комбинаций. Важно отметить, что в образованном множестве комбинаций можно заметить определённые закономерности, которые носят циклический характер. В комбинаторике подобные структуры получили название орбит. Для каждой орбиты можно указать наименьшее число, которое лежит в основе циклического сдвига комбинации из k неповторяющихся элементов (чисел). Назовём такое число образующей комбинацией цикла (ОКЦ). Добавляя номер к ОКЦ, можно указать номер орбиты в общей системе перестановок, которые в последующем целесообразно связать с конкретной порождающей матрицей эквивалентного кода (ЭК), которая лежит в основе системы ПД. Поскольку перестановки представляют группу, то представляется удобным связать эту группу с групповыми кодами, которые формируются над двоичными полями Галуа или их расширениями. Следовательно, разрабатываемый материал системы ПД пригоден как для двоичных кодов, так и для недвоичных избыточных кодов. Пусть дан двоичный группой код Хэмминга (7, 4, 3). Занумеруем элементы кодовой комбинации кода от 1 до $n=7$ и назовём подобные номера нумераторами позиций. В этом случае для рассматриваемого кода допустимо сформировать 35 перестановок, которые удобно разбить на орбиты с их ОКЦ так, как это представлено в таблице 1.

Таблица 1

Орбиты перестановок для кода (7, 4, 3)

1234 ОКЦ¹	1236 ОКЦ²	1245 ОКЦ³	1246 ОКЦ⁴	1235 ОКЦ⁵
2345	2347	2356	2357	2346
3456	1345	3467	1346	3457
4567	2456	1457	2457	1456
1567	3567	1256	1356	2567
1267	1467	2367	2467	1367
1237	1257	1347	1357	1247

Заметно, что в таблице используется правое движение по циклу, которое назовем прямым. В случае использования левого движения по циклу назовем его обратным. Весом любой перестановки является сумма цифр в нее входящих. Тогда становятся очевидными несколько свойств произвольной ОКЦ¹, играющих важную роль в вопросах минимизации объема памяти ККД.

Методы

Свойство 1. Любая ОКЦ имеет минимальное значение веса относительно весов всех иных элементов орбиты, которую она формируют. Это свойство объясняется тем, что именно нумераторы ОКЦ занимают левые позиции в ряду чисел цикла от 1 до k и движение по орбите в прямом направлении хотя бы на шаг до значения n только увеличивает вес перестановки. В случае прямого движения по циклу переход цифры младшего (правого) разряда перестановки равного n в значение $(n+1)$ тождественно операции вида $(n+1) \equiv 1$ с установкой единицы в старший разряд и увеличению цифр остальных разрядов на единицу. Обратная операция в случае обратного движения по циклу при условии, что старший разряд перестановки равен единице равносильна движению всех цифр других разрядов влево с установкой на месте старшего разряда значения $(n+1)$ и уменьшению всех цифр перестановки на единицу.

Свойство 2. В случае полного цикла общее количество перестановок любой орбиты, начинающихся с цифры 1 всегда равно значению k . Это легко доказывается составом элементов орбит таблицы 1.

Установим правило, при котором возможна минимизация веса перестановки. Для минимизации веса перестановки необходимо из цифр всех разрядов вычесть такое число, которое обеспечивает получение в старшем разряде единицы. Например, в перестановке вида 3457 процедура минимизации приводит к перестановке 3457-2222=1235. Указанное правило играет важную роль в ускорении процедуры поиска требуемой ОКЦ для произвольной перестановки.

Свойство 3. Правило минимизация веса перестановки по признаку 1 в левом разряде позволяет реализовать ускоренный поиск ОКЦ и поэтому точно назвать номер орбиты. Пусть из таблицы 1 декодеру в ККД известно содержание только верхней строки и пусть дана произвольная перестановка в лексикографическом формате (ЛФ). Применяя правило обратного цикла и минимизируя перестановку, всегда можно найти перестановку с минимальным весом.

Анализ результатов

Результаты имитационного моделирования системы с ПД. Оценим возможности реализации ПД для некоторых кодов. Пусть код является максимально декодируемым кодом, для которого $d_{\min} = n - k + 1$. Асимптотическая оценка ЭВК для подобных кодов

определяется выражением $D_{\max}(k) = 10 \lg(k - \frac{k^2}{n} + \frac{k}{n})$ [5]. Оценка экстремума указанной

функции приводит к выражению вида $2k = n + 1$. Это означает, что ПД эффективно, когда длина параметра k составляет половину кодового вектора n . Исследование разработанной имитационной модели на основе канала с гауссовой помехой показало эффективность применения выражения (2) для формирования МРС. Результаты испытаний представлены на рисунке 1.

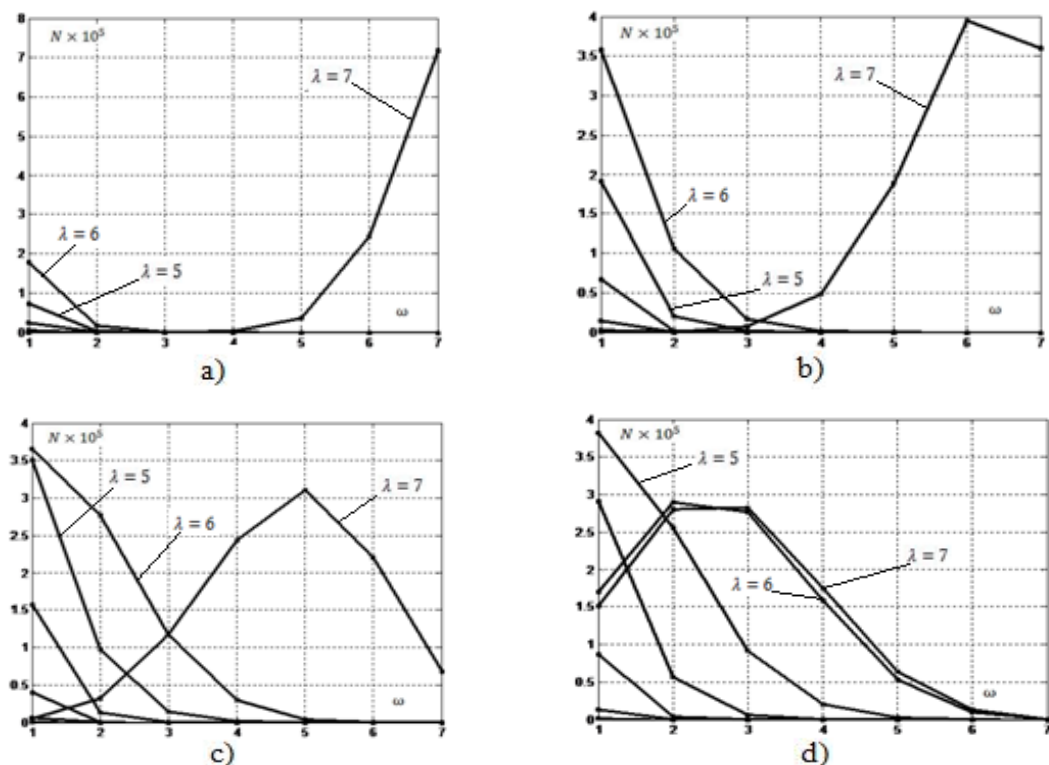


Рис. 1. Частота появления ИМР в кодовом векторе длины $n = 7$, где: а) $\rho = 0.6$; б) $\rho = 0.7$; в) $\rho = 0.8$; д) $\rho = 0.9$

Заметно, что при различных параметрах ρ и отношении сигнал/шум в 10дБ характерного для оптических систем обработка данных в режиме ПД вполне реализуема, поскольку число наиболее надежных МРС превосходит половину длины кодового

вектора. Расширение интервала стирания приводит к уменьшению максимума надежных оценок, но в сумме такие оценки существенно превосходят ближайшие к ним МРС. Так на графике d) оценка с $\lambda=5$ существенно уступает сумме оценок с $\lambda=6$ и $\lambda=7$. Увеличение длины кодового вектора практически не изменяет сделанных выводов. Результаты для кодовых векторов длины 15 показаны на рисунке 2.

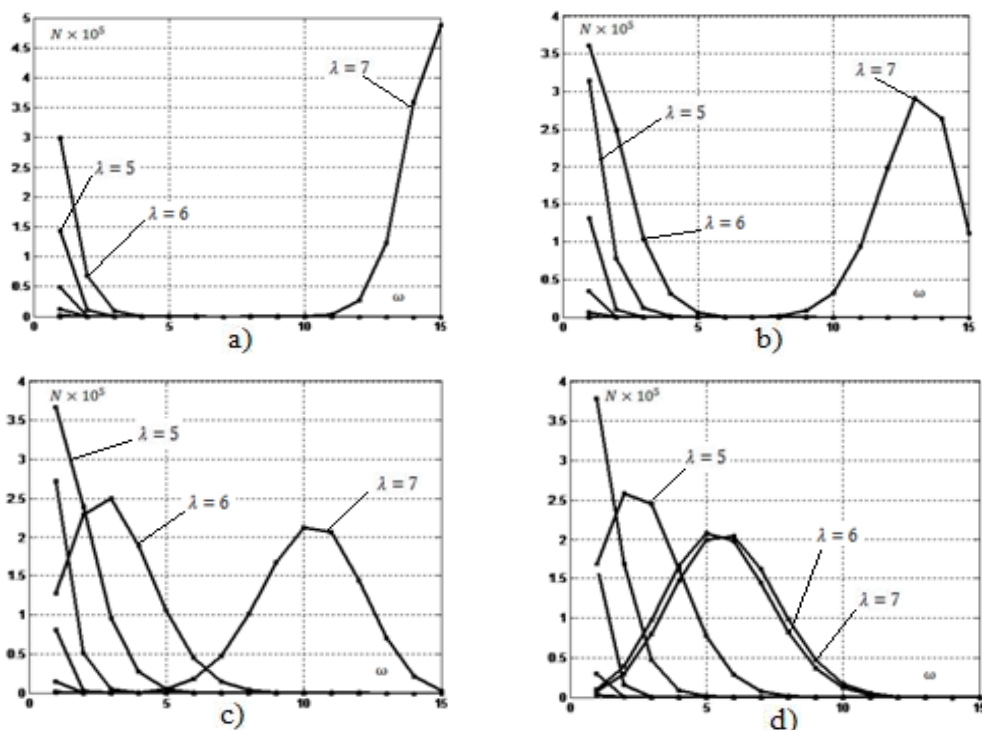


Рис. 2. Частота появления ИМР в кодовом векторе длины $n = 15$, где: а) $\rho = 0.6$; б) $\rho = 0.7$; в) $\rho = 0.8$; д) $\rho = 0.9$

Принципиально стирающий канала связи может быть заменен на систему с полярным кодированием, но оценка подобного решения требует дополнительного исследования.

Заключение

В работе введено новое понятие ОКЦ и на этой основе представляются свойства орбиты, показывается быстрый алгоритм поиска ОКЦ в случае получения декодером произвольной перестановки символов кодового вектора. Связывая конкретную ОКЦ с порождающей матрицей ЭК, легко решается задача поиска соответствующего вектора ошибок, действовавшего в канале связи, на базе уже известных алгоритмов.

Для поиска МРС, лежащих в основе ПД, предлагается использование стирающего канала связи. Указываются параметры линейных функций, обеспечивающих получение значений МРС в целочисленном формате или в формате действительных чисел по решению конструктора приемника.

Приводятся результаты статистических испытаний имитационного моделирования канала связи для оптической системы связи, которые показали достижение полученных границ для числа надежных символов на длине кодового вектора в условиях использования двух практически значимых избыточных кодов.

Применение полярных кодов [7] не противоречит принципам ПД и может быть успешно использовано в таких системах, поскольку систему стертых позиций можно идентифицировать как перестановку из множества ненадежных символов.

Библиографический список

1. Гладких А.А. Основы теории мягкого декодирования избыточных кодов в стирающем канале связи. – Ульяновск: УлГТУ, 2010. – 379 с.
2. Гладких А.А. Перестановочное декодирование как инструмент повышения энергетической эффективности систем обмена данными // Электросвязь. – 2017 - № 8, С. 52–56.
3. Гладких А.А., Пчелин Н.А., Шахтанов С.В. Минимизация объёма памяти когнитивной карты декодера в системе поиска эквивалентных кодов // Радиотехника. – 2018. – № 6. С.38-41.
4. Шахтанов С.В. Перестановочное декодирование на базе упорядоченной когнитивной карты // Сборник докладов 1-й Всероссийской конференции «Современные технологии обработки сигналов» СТОС-2018. Москва, 2018 г. С.190 – 194.
5. Gladkikh A. A, Volkov An. K., Volkov Al. K., Andriyanov N.A., Shakhtanov S.V. Development of network training complexes using fuzzy models and noise-resistant coding Proceedings of the International Conference on Aviaemechanical Engineering and Transport (AviaENT 2019), 2019, pp. 373–379.
6. Gladkikh A. A, Ganin D. V., Pchelin N. A., Shakhtanov S.V., Ochepovsky A. V. Coding Methods and Permutation Decoding in the Systems for Network Processing of Data International Journal of Control and Automation Vol. 13, No.1, (2020), pp. 93 – 110.
7. Linghui Zhon. Polar Codes for Identification Systems, Degree project in electrical engineering, Second Cycle, 30 credits Stockholm, Sweden 2018, p. 55.

S.V. Shakhtanov

REMOTE DECODING BASED ON CYCLIC PROPERTIES OF SEQUENCES OF CODE VECTOR SYMBOLS

Nizhny Novgorod State Engineering and Economic University, Knyaginino

To speed up the process of decoding the received code vector from the communication channel on the basis of permutation decoding, it was previously proposed to use a cognitive map, in the memory of which all variants of combinations of symbols are stored to exclude carrying out matrix calculations when permuting a new code vector. However, in some cases, a large amount of cognitive map memory is not required. It was found that all symbolic sequences coming from the communication channel have a cyclic nature, and the properties of the orbits of cyclic permutations can significantly reduce the memory size of the cognitive map and, accordingly, the decoding time of the code sequence.

Keywords: code vector, permutation decoding, cognitive map, orbits of permutations.

К ВОПРОСУ ОЦЕНКИ СЕТИ СВЯЗИ НА ОСНОВЕ КОГНИТИВНОГО РАДИО

¹АО «Конструкторское бюро «Кунцево», г. Нижний Новгород²Нижегородский государственный технический университет им. Р.Е. Алексеева

В статье рассматривается система когнитивного радио и стандарт, построенный по ее принципу. Затрагиваются вопросы касательно методов рационального использования спектра и способов их реализации. Приводятся общие принципы построения сетей связи по стандарту IEEE802.22, особенности, достоинства и недостатки стандарта.

Ключевые слова: связь, сеть, частотный диапазон, когнитивное радио, радиоканал, стандарт.

Интенсивное развитие беспроводной связи вызывает новую проблему – загруженность радиочастотного спектра. Для работы сотовых телефонов, ЭВМ, подключенных к беспроводной сети, радиостанций и прочим устройствам работать корректно и не создавать помех другу, следует использовать весь радиодиапазон рационально. При этом необходимо, чтобы электронные устройства сами анализировали ситуацию и могли выбирать наиболее подходящий способ и протокол связи. Повсеместное распространение беспроводных сетей связи, таких как системы сотовой и спутниковой радиосвязи, локальные беспроводные сети по технологии Wi-Fi и WiMAX, выявило серьезную проблему. К настоящему времени практически весь радиочастотный диапазон распределен и при этом используется малоэффективно. Как следствие, введение и использование новых сервисов, для которых требуется наличие свободных частотных диапазонов, становится затруднительным или невозможным. Повысить эффективность использования спектра могут комплексы динамического управления спектром, одним из которых является когнитивное радио (КР).

Когнитивное радио позволяет рационализировать использование спектра за счет передачи данных на незанятых в данный момент частотах, не мешая при этом лицензированным пользователям спектра. Когнитивная радиосистема – это радиосистема, в работе которой учитываются сведения об окружающей рабочей и географической средах, об установившихся правилах и ее внутреннем состоянии. В соответствие с этими сведениями, системы в динамическом режиме автономно подстраивают параметры работы и протоколы передачи с тем, чтобы достичь поставленной цели и одновременно учиться на основании полученного опыта. Особенно такой подход важен для повышения живучести сети постоянно перемещающихся мобильных устройств, ведущих непрерывный обмен информацией между собой. Так как такая сеть состоит из устройств (узлов) способных выступать в роле резервных узлов (или ретрансляторов), то показателем живучести сетей связи может быть вероятность того, что сохранится связь с не менее m узлов из n и она может быть рассчитана с использованием частной теоремы для вероятностей появления событий при повторных независимых испытаниях (1) или (2)

$$P_{n,m} = \sum_{k=m}^n C_n^k p^k q^{n-k}; \quad (1)$$

$$P_{n,m} = 1 - \sum_{k=0}^{m-1} C_n^k p^k q^{n-k}. \quad (2)$$

где p -вероятность сохранения узла, $q=1-p$ – вероятность выхода узла из строя. Тогда вероятность того, что сохранится хотя бы один узел, определится как:

$$P_{n,1} = 1 - q^n.$$

Однако такой подход к вычислению показателя живучести применим, если все устройства однотипны с точки зрения живучести. Если узлы отличаются по уровню живучести, то показатель живучести сети следует считать с использованием производящей функции:

$$\varphi_n(z) = \prod_{i=1}^n (q_i + p_i z) = \prod_{j=0}^n a_j z^j.$$

p_i – вероятность сохранения i -го узла при заданной модели нанесения удара; q_i – вероятность выхода из строя i -го узла; $\varphi_n(z)$ – производящая функция n -го порядка; z – формальный параметр; a_i – вероятность того, что сохранится ровно i узлов, например:

$$a_0 = \prod_{i=1}^n q_i$$

$$a_n = \prod_{i=1}^n p_i$$

Тогда вероятность сохранения не менее m узлов из n имеющихся:

$$P_{n,m} = 1 - \sum_{j=m}^n a_j$$

Когнитивная радиостанция функционирует согласно когнитивному циклу, предложенному Митолой [1, 2]. Свойство когнитивности (самообучения) подразумевает способность радиосистемы решать такие задачи как:

- Оценка так называемой шумовой температуры радио-среды, обнаружение неиспользуемых в данный момент времени спектральных диапазонов. Задача состоит в том, чтобы непрерывно вести мониторинг спектра на наличие не занятых частотных диапазонов (в литературе могут встречается следующие названия: спектральные дыры, белые пятна, пробелы) для дальнейшего использования.

- Анализ параметров радиоканала, оценка канальной информации, предсказание состояния радиоканала. Система должна понимать на что способен канал и иметь понятие о его характеристиках, а также должна понимать какая информация передается, например, какие объёмы и уметь спрогнозировать состояние канала, чтобы в экстренный момент перебросить поток данных на другие диапазоны частот, подходящие для этого.

- Контроль излучаемой мощности и динамическое управление спектром. Задача контроля излучаемой мощности в многопользовательской среде сводится к выбору оптимального уровня излучаемой мощности среди множества числа абонентов с целью максимизации совокупной скорости передачи, не превышая максимально допустимого уровня шумовой температуры и при условии ограниченного числа спектральных дыр. Проблема заключается в том, что при увеличении мощности передачи одного из пользователей наблюдается нежелательный эффект увеличения уровня интерференции на входе приемников других пользователей. Поэтому необходимо соблюдать баланс между излучаемой мощностью и уровнем создаваемых помех для других абонентов.

На основе когнитивного радио построен стандарт беспроводных региональных сетей IEEE 802.22 [3], в нем для передачи данных используются три варианта многопозиционной модуляции (QPSK, QAM-16, QAM-64), четыре возможных скорости помехоустойчивого кодирования (1/2, 2/3, 3/4, 5/6), четыре различных длины циклического префикса: 1/4, 1/8, 1/16, и 1/32 от длительности информационного символа.

Такое разнообразие параметров позволяет выбрать наилучшее соотношение скорости передачи данных и надежностью сети связи в текущей обстановке [4]. Сеть может работать как фиксированными базовыми станциями (БС), так и с портативными пользовательскими терминалами, однако для большей зоны покрытия необходимо использовать больше мощности излучения. Для этих целей предусмотрено размещение изотропных или секторных антенн на БС, в зависимости от расположения абонентов. На клиентской стороне используется два вида антенн, узконаправленная и сканирующая. Первая, ориентируется в сторону БС или БС с наилучшим сигналом (если БС несколько). Вторая антенна используется непосредственно для функционирования когнитивных механизмов — это всенаправленная антенна, обеспечивающая чувствительность к вертикальной и горизонтальной поляризации, что позволяет воспринимать сигнал от телевизионных станций и беспроводных устройств. При необходимости к двум имеющимся антеннам может быть размещена GPS-антенна если используется спутниковое позиционирование. Существенным недостатком стандарта IEEE 802.22 является невозможность поддержки технологий множественных антенн (MIMO). Это связано с тем, что в рабочем диапазоне частот физически сложно обеспечить достаточное разнесение нескольких антенн, которое должно составлять не менее трех длин волн.

Использование адаптивных направленных когнитивных антенн составе системы КР позволяет получить такие преимущества как [5]: увеличение дальности связи и пропускной способности, вероятность своевременной доставки пакета, помехозащищенность, уменьшение задержки передачи пакета. Когнитивная антенна способна распознавать параметры приходящей помехи, она имеет адаптивную диаграмму направленности, что позволяет менять направление и ширину луча, однако это свойство подлежит тщательному контролю. Так, например, если луч антенны широкий, то вероятность приема при движении выше, пропускная способность больше, однако это также повысит внутрисистемные помехи, снизит отношение сигнал/шум, помехозащищенность и пространство для поиска параметров.

Заключение

Когнитивное радио развивающаяся технология, у которой в современных реалиях есть большой потенциал. Применение технологии когнитивное радио позволяет эффективно использовать имеющийся спектр, что влечет за собой не только повышение живучести сети связи, но и уровня предоставляемых услуг.

Преимущество модели использования радиочастотного спектра систем когнитивного радио так же является его недостатком, так как оно не гарантирует наличие свободных частотных диапазонов, которые могут быть использованы без влияния на соседние в течение времени эксплуатации сети. Работа системы когнитивного радио описана в стандарте IEEE 802.22 и отличие этого стандарта от других в том, что он рассчитан на работу в сельской местности или регионах, где частотный спектр не столь загружен и базовые станции имеют гораздо большую область покрытия, чем, например, в городе. В связи с этим требуются дополнительные исследования для расширения возможностей использования когнитивного радио.

Библиографический список

1. Mitola, J. Cognitive radio architecture: the engineering foundations of radio XML. New Jersey: Wiley&SonsInc. Publication, 2006. 488 p.
2. Dong, X., Li Y., Wei S. Q. Design and implementation of a cognitive engine functional architecture // Chin Sci Bull. 2012. Vol.57. P. 3698-3704.
3. IEEE 802.22 Working Group on Wireless Regional Area Networks. – URL: <https://www.ieee802.org/22/> (датаобращения 20.03.2021). - Текст: электронный.

4. Региональное содружество в области связи. Справочник по перспективам внедрения систем когнитивного радио в диапазоне УВЧ в странах участников РСС. /Астана, 2016 г.- 51с.

5. Feng, O., Jerry H., Henry S., Robert N. Robust cognitive networks interfacing with cognitive antennas//MILCOM 2010 Conference Paper. December 2010.P. 1689–1694.

A.V. Pronin^{1,2}

**ON THE ISSUE OF ASSESSING COMMUNICATION NETWORKS BASED ON
COGNITIVE RADIO**

¹JSC "Design Bureau" Kuntsevo ", Nizhny Novgorod

²Nizhny Novgorod state technical university n.a. R.E. Alekseeva

The article discusses the cognitive radio system and the constant standard according to its principle. The questions are raised concerning the methods of rational use of the spectrum and the ways of their implementation. The general principles of building communication networks according to the IEEE 802.22 standard, features, advantages and disadvantages of the standard are given.

Keywords: communication, network, frequency range, cognitive radio, radio channel, standard.

ОПТИМИЗАЦИЯ СИГНАЛА ДЛЯ ИСПЫТАНИЙ СИСТЕМ И КОНСТРУКЦИЙ

*Нижегородский технический университет им. Р.Е. Алексеева,
Высшая школа экономики, Нижний Новгород*

В работе был проведен анализ недостатков традиционных способов испытаний систем и конструкций, к числу которых относится отсутствие согласованности характеристик испытательного сигнала и объекта испытаний. Это не позволяет получить при прочих равных условиях максимально возможную реакцию объекта испытаний на испытательный сигнал. Поэтому в условиях эксплуатации объект, прошедший испытания, может быть разрушен. Указаны причины несовершенства традиционного подхода при формировании испытательного сигнала и решена задача оценки величины проигрыша его интенсивности при использовании традиционного способа испытаний. Показана возможная величина коррекции интенсивности испытательного сигнала в традиционном способе испытаний при условии дополнительного измерения фазо-частотной характеристики объекта испытаний.

Ключевые слова: испытательный сигнал, шумовой способ испытаний, максимальная реакция объекта испытаний, согласование характеристик испытательного сигнала и объекта испытаний.

Постановка задачи

Существующие способы испытаний конструкций и систем отличаются многообразием принципов формирования испытательного сигнала. Некоторые из них заключаются в воздействии на объект испытаний заданными случайными широкополосными вибрационными, ударными и гармоническими нагрузками. Вибрационные испытания проводят по методу качающейся частоты, когда частоту вибрации плавно изменяют в заданном диапазоне частот от нижней границы частоты к верхней и обратно при постоянстве заданных параметров вибрации в течение определенного времени. Испытания проводят также по методу случайной широкополосной вибрации, когда одновременно возбуждают все резонансные частоты объекта. За критерий подобия реальному процессу принята спектральная плотность мощности виброускорений. Ударные испытания проводят по методу ударных спектров ускорений, когда не важен вид воздействия, а важна реакция на это воздействие. В ряде случаев динамические испытания проводят в комбинированном режиме. Во-первых, определяют резонансные частоты объекта испытаний во всем нормируемом частотном диапазоне и устанавливают частотные диапазоны, в которых проводят замену случайной широкополосной вибрации эквивалентной ей гармонической вибрацией и в процессе испытаний проводят соответствующие измерения. Однако следует указать и на недостатки этих способов испытаний.

Недостатком этих способов испытаний является то, что одна или несколько собственных и резонансных частот, заключенных в широкополосном спектре сигнала воздействия, на которых проводятся испытания, а фазовые соотношения которых являются произвольными, не отражают всех особенностей комплексной амплитудно-частотной характеристики конструкции или системы как объекта испытаний. И даже при охвате всего диапазона частот амплитудно-частотной характеристики объекта вышеуказанными испытательными сигналами не достигается полного согласования комплексного спектра испытательного сигнала с параметрами комплексного коэффициента передачи конструкции. Это происходит вследствие того, что поскольку широкополосные вибрационные и ударные нагрузки, используемые в известном способе

испытаний, представляют собой случайные процессы, то, как амплитуды, так и фазы составляющих спектра воздействия никак не коррелированы с соответствующими как амплитудами, так и фазами составляющих амплитудно-частотной характеристики испытуемого объекта. Таким образом, случайный характер спектра диагностического шумового воздействия не соответствует этим требованиям, так как воздействие носит случайный характер. И не позволяет решить главную задачу, задачу получения при прочих равных условиях максимально возможной реакции объекта испытаний на испытательный сигнал. В связи с этим назрела потребность в новых способах, не содержащих указанного недостатка.

Этот факт был использован при исследованиях так называемого резонанса формы [1-3], и основанного на нём способа испытаний [4,7], способа формирования испытательного сигнала [5-6]. Предложенный при этом способ получения максимально возможной реакции объекта при прочих равных условиях на испытательный сигнал основан на успешном решении задачи полного согласования его амплитудно-частотных параметров с соответствующими характеристиками объекта испытаний [8].

Казалось бы, предложенное решение проблемы должно вызвать интерес специалистов в области испытаний. Однако, в силу ряда причин, в том числе нездоровый консерватизм специалистов в области испытаний, помешал его немедленному использованию. Это, в свою очередь, стимулировало авторов нового способа испытаний показать сравнительную оценку преимуществ нового способа и величину проигрыша в эффективности традиционного способа испытаний. В связи с этим появилась идея о возможности коррекции характеристик испытательного сигнала в существующих традиционных способах испытаний для повышения их эффективности. Испытательный сигнал при этом, при прочих равных условиях сможет обеспечить максимальную реакцию объекта и исключить его разрушение в реальных условиях эксплуатации с учетом предложенных подходов к решению поставленной задачи.

Решение задачи

Поставленная задача решается с помощью доказательства существования способа численного определения максимально возможной при прочих равных условиях величины реакции объекта испытаний за счёт соответствующей коррекции величины традиционного испытательного сигнала. Это доказательство возможно при использовании наряду с АЧХ объекта дополнительно его фазово-частотной характеристики, т.е. в итоге при использовании комплексного коэффициента передачи объекта испытаний.

Доказательство

Воздействие на систему описывается гармониками комплексного ряда Фурье с коэффициентами (размерность В) с амплитудами:

$$\{U_m\}_{m = M, -M + 1, \dots, -1, 0, 1, \dots, M}$$

Система описывается комплексной частотной характеристикой $\{G_m\}$, представляющей собой конечное (счётное) множество комплексных коэффициентов (безразмерных) на каждой из дискретных частот m .

$$\{G_m\}_{m = -M, -M + 1, \dots, -1, 0, 1, \dots, M}$$

Отклик системы может быть представлен как комплексный ряд Фурье с коэффициентами (с физической размерностью В):

$$\{S_m = U_m G_m\}_{m = -M, -M + 1, \dots, -1, 0, 1, \dots, M}$$

Максимальное значение отклика системы во временной области на вышеуказанное произвольное воздействие в некоторый момент времени можно записать так:

$$s^{max} = \sum_m |U_m G_m| \cos \varphi_m$$

где φ_m - фазы частотных составляющих комплексного коэффициента передачи системы.

Наиболее возможный максимальный отклик системы s^{opt} может быть получен в соответствии с теорией согласованного фильтра, в роли которого рассматривается объект испытаний, а испытательный сигнал имеет вид зеркального отображения импульсной характеристики объекта. В свою очередь, комплексный спектр импульсной характеристики с точностью до постоянного множителя представляет собой комплексный коэффициент передачи объекта $\{G_m\}$. Зеркальность импульсной характеристики порождает сопряжённость комплексного коэффициента передачи объекта испытаний G_m^* , амплитуды комплексного ряда Фурье которого находятся так:

$$U_m^{opt} = A G_m^* \text{ где } m = -M, -M+1, \dots, -1, 0, 1, \dots, M,$$

где знак * - комплексное сопряжение, А - произвольный коэффициент (размерности В).

Максимально возможная реакция на вышеуказанное оптимальное воздействие определяется как сумма

$$s^{opt} = \sum_m A G_m^* G_m = A \sum_m |G_m|^2$$

Отсюда следует, что максимальная реакция объекта на испытательное воздействие определяется только модулем $|G_m|$ спектральной плотности воздействия и не зависит от его фазо-частотного спектра φ_m . Это является следствием того, что взаимные фазовые сдвиги спектральных составляющих испытательного сигнала компенсируются фазо-частотной характеристикой согласованного с этим воздействием объекта испытаний. Поэтому все спектральные составляющие оптимального воздействия достигают амплитудных значений в момент времени $t = T$ и, складываясь, дают пик реакции объекта испытаний.

Энергия произвольного входного воздействия на интервале времени T может быть представлена выражением:

$$W = \sum_m |U_m|^2$$

Энергия воздействия с большим максимальным откликом имеет вид:

$$W^{opt} = \sum_m |A G_m|^2 = A^2 \sum_m |G_m|^2$$

При этом максимальная реакция системы на испытательное воздействие примет некоторую вышеуказанную величину s^{opt} .

Приравняем энергии произвольного воздействия и энергии воздействия, дающего больший максимальный отклик.

$$W^{opt} = W = \sum_m |U_m|^2 = A^2 \sum_m |G_m|^2$$

Из этого равенства может быть определён коэффициент А при заданной энергии входного воздействия

$$A = \sqrt{\frac{\sum_m |U_m|^2}{\sum_m |G_m|^2}}$$

Теперь определим, во сколько раз максимальный отклик системы при воздействии согласованного сигнала превышает максимальный отклик при воздействии произвольного сигнала:

$$k = \frac{\sqrt{\sum_m |U_m|^2 \sum_m |G_m|^2}}{\sum_m |U_m G_m| \cos \varphi_m}$$

Коэффициент k позволяет при произвольном (не согласованном) испытательном воздействии измерить максимальную величину реакции на это произвольное воздействие, умножить эту величину на коэффициент k , а затем увеличить величину произвольного воздействия так, чтобы максимальная реакция системы достигла величины S^{opt} , т.е. величины максимальной реакции системы, если бы она находилась под согласованным воздействием. При этом величина реакции системы будет превышать результат, получаемый при традиционном способе испытаний, что в значительной степени приближает эффективность скорректированного традиционного способа к максимально возможной эффективности, достигаемой при реализации согласованной формы испытательного сигнала, соответствующей зеркальному отображению импульсной характеристики объекта испытаний.

Заключение

Полученные в работе результаты свидетельствуют о том, что традиционные способы испытаний систем и конструкций, в которых не уделено должного внимания характеристикам испытательного сигнала в смысле их соответствия характеристикам объекта испытаний, тем не менее могут быть избавлены частично от этого недостатка за счет обоснованной коррекции величины испытательного сигнала, позволяющей избавиться от возможных рисков разрушения объекта в реальных условиях эксплуатации.

Библиографический список

1. Зельманов С.С. Исследование явления резонанса формы сигнала в согласованном фильтре / С.С. Зельманов // Электросвязь. 2011, №1. С. 34-38.
2. Зельманов С.С., Крылов, В.В. О роли резонанса формы при диагностике состояния систем и конструкций. / С.С. Зельманов, В.В. Крылов // Т•Сотт (Телекоммуникации и транспорт), №5, 2014г., С.12-15.
3. Зельманов С.С., Крылов, В.В. Диагностика состояний объектов и процедуры их испытаний на основе резонанса формы. // Известия ЮФУ. Технические науки. Тематический выпуск «Компьютерные и информационные технологии в науке, инженерии и управлении». Раздел 2, Методы, модели и алгоритмы обработки информации. - Таганрог: Изд-во ТТИ ЮФУ, 2015г. №3, С. 23-28.
4. Патент RU № 2569636 С, кл. G01M/ (2006.01). Способ динамических испытаний конструкций и систем на механические и электронные воздействия // С.С.Зельманов, В.В.Крылов, заявка 2014108466/28, 04.03.2014, заявлено 04.03.2014, опубликовано 27.11.2015 Бюлл. № 33.
5. Зельманов С.С. Крылов В.В. Формирование оптимального диагностического сигнала при испытаниях систем и конструкций / Зельманов, С.С. Крылов В.В./ XII Международная научная конференция «Перспективные технологии в средствах передачи информации – ПТСПИ-2017». Владимирский государственный университет имени А.Г и Н.Г. Столетовых. Суздаль, Россия, июль 2017, т.1, С. 102-105.
6. Зельманов С.С., Крылов, В.В. Оценка диагностических сигналов для испытаний систем и конструкций / С.С.Зельманов // Датчики и системы. №5 2018, С. 10.
7. Патент RU № 2702930 С, кл. G01M/ (2006.01). Способ динамических испытаний конструкций и систем на механические и электронные воздействия // С.С.Зельманов, В.В.Крылов, заявка № 2019107896, заявлено 19.03.2019. Дата госрегистрации в госреестре РФ 14.10.2019. Срок действия исключительного права 19.03.2039г.
8. Зельманов С.С., Крылов, В.В., Пятериков В.И. Диагностирование потенциально максимальных деформаций при динамических испытаниях синтезированными нагрузками/ С.С. Зельманов, В.В. Крылов, В.И. Пятериков // Датчики и системы, №10, 2019, С.8.

S.S. Zelmanov, V.V. Krylov

ONLINE SIGNAL OPTIMIZATION FOR TESTING SYSTEMS AND STRUCTURES

*Nizhny Novgorod State Technical University n.a. R. E. Alekseev
High School of Economics, Nizhny Novgorod*

The paper analyzes the shortcomings of traditional methods of testing systems and structures, which include the lack of consistency of the characteristics of the test signal and the test object, this does not allow us to obtain, all other things being equal, the maximum possible reaction of the test object to the test signal. Therefore, under operating conditions, an object that has passed the test can be destroyed. The reasons for the imperfection of the traditional approach in the formation of the test signal are indicated and the problem of estimating the magnitude of the loss of its intensity when using the traditional test method is solved. The possible value of correction of the intensity of the test signal in the traditional test method is shown, provided that the phase-frequency characteristic of the test object is additionally measured.

Keywords: test signal, noise test method, maximum response of the test object, matching the characteristics of the test signal and the test object.

ИССЛЕДОВАНИЕ ПОКАЗАТЕЛЕЙ СВЯЗАННОСТИ СЕТИ БЕСПИЛИТНЫХ ЛЕТАТЕЛЬНЫХ АППАРАТОВ

¹ НПП «Прима», Нижний Новгород

² Нижегородский государственный технический университет им. Р.Е. Алексеева

Проведено исследование свойств связанности сети FANET в условиях нестабильности на канальном уровне. Разработан алгоритм для исследования связанности быстроменяющейся конфигурации сети на основе теории графов. Проведено моделирование с помощью симулятора NS-3. Показано, что возможно достижение достаточно высоких показателей потенциальной и реальной связности путем внесения интеллектуального контроля за движением БПЛА.

Ключевые слова: сеть FANET беспилотных летательных аппаратов (БПЛА), связанность сети, узлы сети.

Применение беспилотных летательных аппаратов (БПЛА) находит широкое распространение для самых различных целей. Для решения некоторых задач актуально применение множества БПЛА, образующих подвижную самоорганизующуюся сеть. Высокие скорости полета, большой объем расположения в пространстве, меняющиеся условия каналов передачи, множественность траекторий перемещения, удаленность от центра управления делает чрезвычайно актуальной проблемы самоорганизации и обеспечения связанности сети, особенно в условиях быстроизменяющейся сетевой топологии.

В данной работе предлагается способ оценки работы протоколов маршрутизации в сетях FANET на основе критерия связности. Разработан алгоритм определения связности сети и проведено моделирование для оценки работоспособности разработанного алгоритма. В работе выбрана модель трафика, при которой один узел сети является источником потоков данных реального времени, передаваемых каждому узлу сети. Потоки являются независимыми и некоррелированными и передаются с помощью UDP дейтаграмм постоянного размера и скорости. пакеты этих потоков трафика будут также служить для расчета таких сетевых характеристик как реальная связность и коэффициент доставки пакетов полезных данных.

Типичным способом задания математической модели сетей связи является использование графов. Идея заключается в построении графа $G = (V(G), E(G))$ некоторой сети, в которой каждый маршрутизатор (БПЛА) соответствует вершине графа из множества $V(G)$, а каналы связи между БПЛА — множеству ребер $E(G)$ графа. Два узла v и u считаются соединенными, т.е. являются концами некоторого ребра e , если они могут связываться напрямую в соответствии с правилами установления каналов связи и доступа к общей среде распространения. Такие узлы сети являются смежными.

Таким образом, модель сети представляет собой: узел (БПЛА) – вершина графа, радиоканал между двумя БПЛА – ребро графа. Т.е. в модели сети два узла соединены ребром в том случае, если они находятся в зоне уверенного приема сигнала друг от друга, а также возможно установление канала связи между этими узлами в соответствии с технологией MAC-уровня. Следовательно, такой граф будет описывать логическую топологию сети в какой-либо момент времени, т.к. при ее описании ребро между вершинами проводится только при условии наличия канала связи, установление которого напрямую зависит от способа организации доступа узлов к разделяемой среде.

Потоки информации между произвольными узлами сети можно также представить графом H , отличие которого, заключается в том, что он отражает потоки информации в

рамках логической топологии. Он показывает пути следования пакетов с данными между узлами сети. Т.е. путь – последовательность вершин $v_1 v_2 \dots v_n$ и ребер $e_1 e_2 \dots e_{n-1}$ в графе, что так же можно назвать маршрутом. Отысканием и построением этих путей в рамках графа сетевого уровня как раз и занимается алгоритм маршрутизации.

Это означает, что граф H на уровне сетевого взаимодействия состоит из тех же вершин $V(H) = V(G)$ графа G уровня передачи данных, однако ребра этого графа не всегда в точности совпадают. Таким образом, граф H является остовным подграфом G , индуцированного на всем множестве вершин $U \in V(G)$ исходного графа. Не совпадение может быть обусловлено особенностью работы алгоритма маршрутизации, который устраняет пути с одинаковыми концами u и v , но различными длинами, а также устраняем цикличность в обходе вершин графа. Таким образом, получается направленный ациклический граф. Результат работы протокола маршрутизации можно описать множеством остовных деревьев H , индуцированных на всем множестве вершин исходного графа G , описывающего сеть БПЛА на уровне передачи данных, т.е.

$$V(H_i) = V(G), i = \overline{1, v(G)},$$

где $v(G)$ – количество ребер (узлов) в исходном графе.

На рисунке 1 показан подграф исходного графа сети, который является остовным деревом для узла n_7 , т.е. комбинацией путей от узла ко всем остальным.

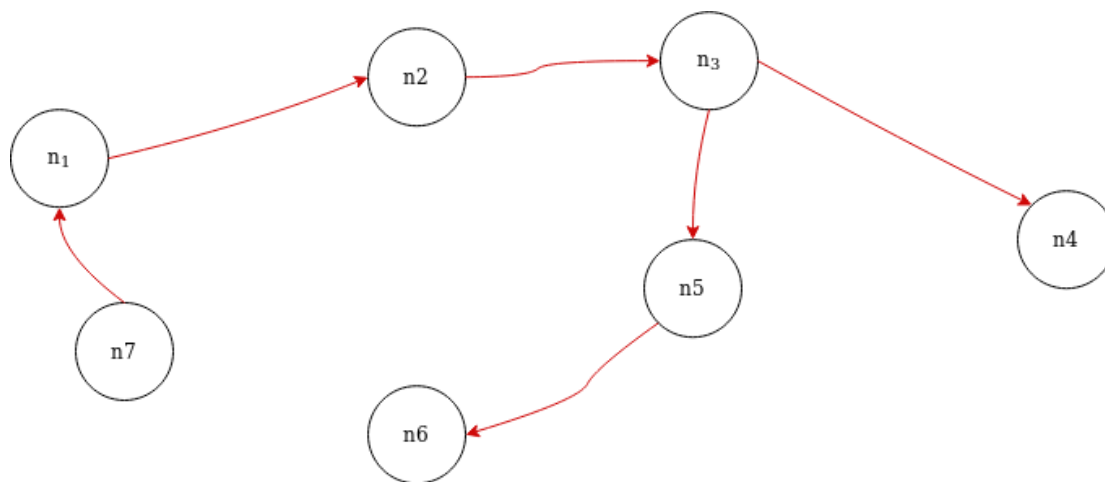


Рис. 1. Остовное дерево для узла n_7 в исходном графе

Таким образом, поведение узлов сети и сети передачи данных можно отображать одинаковым графом логической топологии. Этот граф можно назвать теоретическим, потенциально возможным графом, на который опирается сетевой уровень модели OSI. Данный граф опишет все возможные каналы связи, их направления, дополнительные метрики. На сетевом уровне строится другой граф, являющийся подграфом первого. Этот подграф как раз можно вывести на основе работы протокола маршрутизации, задача которого тем или иным способом выстроить эффективное использование существующих каналов. Данный граф, который будет являться графом информационной топологии, можно назвать реальным, фактическим.

В случае, если для какого-либо узла протоколу маршрутизации не удастся установить маршрут (он изначально не может быть установлен, либо же произошла ошибка в процессе пересылки), то нарушается достижимость между двумя произвольными узлами сети u и v , что означает отсутствие пути между ними. Граф H , описывающий такую сеть, становится несвязным, порождая компоненты связности. Такой пример показан на рис. 2.

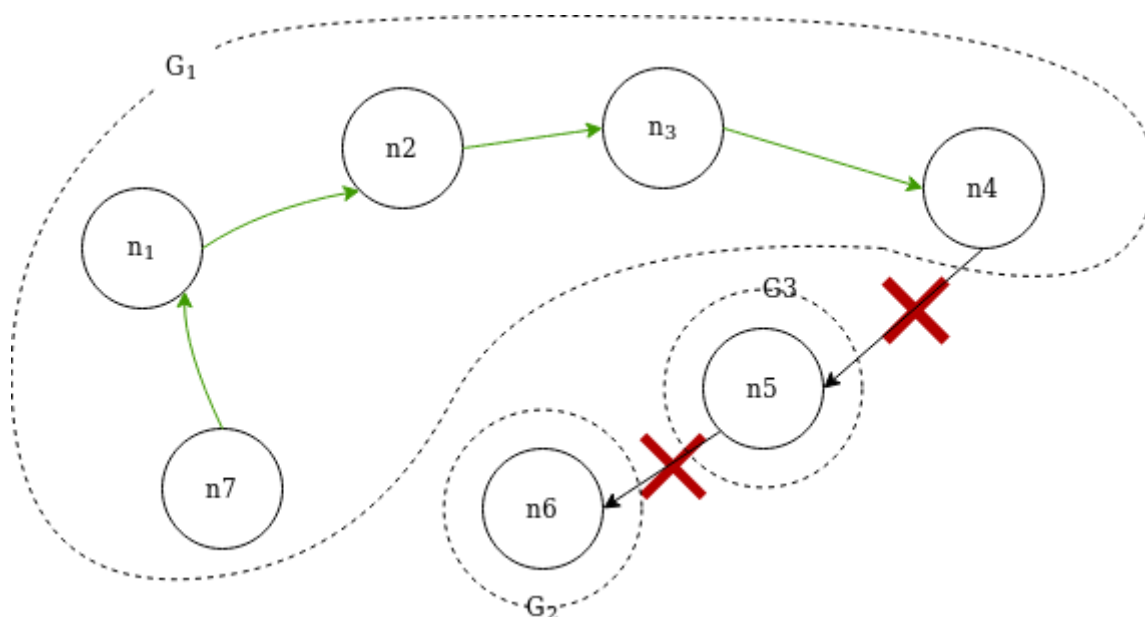


Рис. 2. Нарушение связности графа

Несвязность графа порождает компоненты связности. Компонента связности – это подграф G_i в графе H , который не связан с другими подграфами этого графа.

В изучаемой сети FANET образование более чем одной компоненты связности означает нарушение связности всей сети, т.е. достижимости пакетов от одного узла к другому. В случае, если логическая топология изменяет свой вид таким образом, что потенциальный граф сети становится несвязным, то это неминуемо влечет за собой несвязность и на сетевом уровне. Если узел перемещается в недоступную зону, т.е. нарушается канал связи между двумя смежными узлами, однако граф остается связным, то на это событие, используемый сетевой протокол должен каким-либо образом среагировать, отчего реальный граф сети также изменится. Наиболее удобной средой моделирования является симулятор NS-3, который позволяет эффективно моделировать весьма сложные процессы для подвижных сетей [1].

Логично ввести показатель, который позволит оценивать связность в некотором промежутке времени. Ниже приводится такой алгоритм, который в дальнейшем используется в экспериментальной части.

Если обозначить связность графа через c , то это приводит нас к дискретной функции времени $c(t)$, которая может принимать два значения:

$$c(t) = \begin{cases} 0, \text{ граф несвязный} \\ 1, \text{ граф связный} \end{cases}$$

Предположим, что в произвольный момент времени t есть возможность узнать факт связности графа, т.е. получить значение функции $c(t)$. Если проводить измерение связности достаточно часто, например, через интервалы времени τ_i , то можно получить вектор \vec{c} значений функции связности $c(t)$ в эти отсчеты времени.

Выбор интервала времени τ нужно проводить исходя из данных о том, как часто могут нарушаться каналы связи между узлами сети. Прямое влияние на состояние этих каналов имеет мобильность узлов. В первую очередь здесь можно говорить о средней скорости движения узлов сети в пространстве. Следовательно, можно сделать вывод, что чем больше средняя скорость движения узлов, тем меньше должен выбираться интервал измерений τ связности сети.

$$\tau \sim \frac{1}{v}$$

Если в момент времени τ_i значение связности $c(\tau_i) = 1$, то можно грубо говорить о том, что сеть в интервале времени $(\tau_{i-1}; \tau_i]$ была связна. Тогда время, в течении которого

сеть находилась в связанном состоянии на протяжении нескольких последовательных измерений может быть найдена:

$$t_c = \sum_{i=1}^N \tau_i c_i$$

При периодичном интервале времени между измерениями, когда $\tau = t_1 = t_2 = \dots = t_N$ имеем:

$$t_c = \sum_{i=1}^N \tau_i c_i = \tau \sum_{i=1}^N c_i$$

Тогда на момент времени T симуляции сети можно узнать долю времени, в течении которого сеть была связана:

$$C = \frac{\tau \sum_{i=1}^N c_i}{T}$$

Данный коэффициент и может служить количественным показателем связности сети на некотором интервале времени T . В идеальном случае, когда сеть на протяжении всего наблюдаемого времени находилась в связанном состоянии данный показатель равен 1. В противном случае, показатель связности будет меньше 1. Используя данный критерий, можно проводить сравнения алгоритмов маршрутизации с различными параметрами в рамках одной модели мобильности.

Тесную связь выведенный параметр имеет с такой характеристикой сети как коэффициент доставки пакетов данных (PDR – packet delivery rate). Он показывает, какая доля сгенерированных и отправленных оригинальных пакетов данных было получено узлом назначения за некоторый промежуток времени. Очевидно, что если хотя бы один пакет был потерян, не дошел до адресата, то можно предположить, что реальный граф сети не связан. Таким образом:

$$c(\tau_i) = \begin{cases} 0, & PDR_i < 1 \\ 1, & PDR_i = 1 \end{cases}$$

В случае, когда PDR рассчитывается с помощью отправки по одному пакету каждому узлу в интервале времени, то общее количество полученных пакетов покажет, с каким количеством приемных узлов источник имеет пути. Это можно расценивать как степень связности.

Разработанный в данной работе алгоритм предлагает способ, который бы учитывал связность узлов на канальном уровне, а значит, учитывал бы и сам факт движения узлов.

Степень связности служит хорошим показателем при оценке работы сетевых протоколов и протоколов маршрутизации в рамках одной модели подвижности узлов.

Сравнивая связности реального и потенциального графов можно судить об эффективности E_r работы протокола маршрутизации в текущих условиях мобильности БПЛА. Это можно сделать с помощью отношения показателя связности C_r фактического графа к показателю связности потенциального C_p :

$$E_r = \frac{C_r}{C_p}$$

В данном исследовании в дальнейшем в качестве метода сбора статистических данных используется **метод повторений**. Это обусловлено тем, что это дает больший контроль над последовательными прогонами модели сторонними программными инструментами.

Моделирование движения узлов по всем трем пространственным координатам является одним из наиболее важных моментов при исследовании сетей БПЛА в силу их природы. Для устранения этого недостатка существующая двумерная модель движения GM в сетевом имитаторе ns-3 была доработана до трехмерной в соответствии с моделью, описанной в работе [2]. Данная версия модели хорошо подходит для описания движения летательных аппаратов, где ключевыми параметрами являются направление движения, скорость, угол подъема. Область полета ограничивается параллелепипедом.

В реализованной модели учитывался также угол наклона ЛА (тангаж) и составляющие скорости по трем пространственным координатам.

Проведенные исследования показали эффективность разработанного алгоритма и позволили получать многие характеристики процессов передачи информации в самоорганизующейся сети БПЛА. В результате моделирования получены многие характеристики взаимосвязанной сети. На рис. 3 показана зависимость потенциальной и реальной связности при увеличении количества узлов в сети.

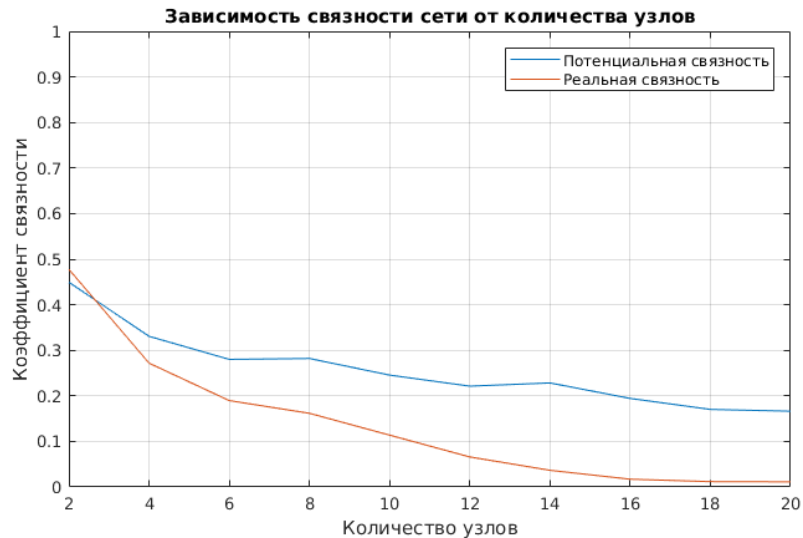


Рис. 3. Зависимости реальной и потенциальной связности от количества узлов

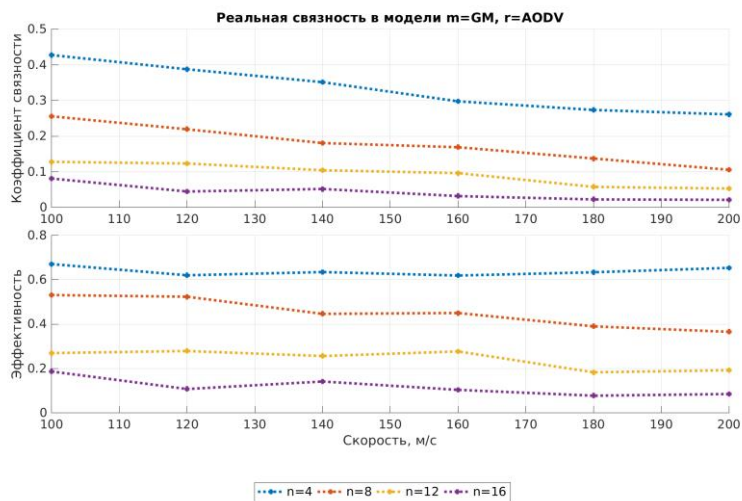


Рис. 4. Показатели реальной связности в зависимости от скорости движения БПЛА

На рис. 4 приведены показатели реальной связности в зависимости от скорости движения ЛА.

Выводы

По результатам данной серии экспериментов для модели мобильности PPRZ можно считать, что траектории движения подобного рода, используемые в поисковых и спасательных операциях, обладают более высокой предсказуемостью, что положительно сказывается на функционировании протоколов маршрутизации в данных условиях. Можно предположить, что в модели PPRZ возможно достижение достаточно высоких показателей потенциальной и реальной связностей путем внесения интеллектуального контроля за движением БПЛА и добавлением корреляции между движениями независимых узлов.

Библиографический список

1. Random Variables [В Интернете] // NS-3: A Discrete-Event Network Simulator. - ns-3 project, 2019 г. - май 2020 г. - <https://www.nsnam.org/docs/release/3.30/manual/html/random-variables.html>.
2. Dan Broyles Abdul Jabbar, and James P.G. Sterbenz Design and Analysis of a 3-D Gauss-Markov Mobility Model for Highly Dynamic Airborne Networks [Конференция] // International Telemetering Conference (ITC 2010). - The University of Kansas Lawrence, KS 66045 : [б.н.], 2010.

А.Е. Ignatev¹, V.S. Syuvatkin²

RESEARCH INDICATORS OF CONNECTIVITY OF THE NETWORK OF UNMANNED AIRCRAFT

¹ *ООО NPP Prima*

² *Nizhny Novgorod State Technical University n.a. R.E. Alekseev*

In this paper, we research the connectivity metric behavior in FANET under instability of data link connections. We develop the algorithm for calculation network connectivity metric for highly changing network configuration using graph theory. The efficiency was analyzed using simulation of test scenarios in Network Simulator (NS-3). To conclude in order to achieve the high rates of network connectivity there is necessary to introduce intelligent UAV movement control as well as highly optimized routing algorithms.

Keywords: FANET network Unmanned aircraft, connectivity, node network.

ОБЗОР ПРОГРАММНЫХ ПРОДУКТОВ ДЛЯ ПОДДЕРЖКИ РАСПРЕДЕЛЕННЫХ КОМАНД

Нижегородский государственный технический университет им. Р.Е. Алексеева

В статье рассмотрена поддержка формирования и функционирования виртуальных команд фрилансеров. Обозначены особенности требуемого программного обеспечения и его классификация. Намечены направления развития этих программных продуктов.

Ключевые слова: виртуальные организации, оценка компетенций, системы поддержки принятия решений, тестирование сотрудников.

Развитие информационных технологий, а также вызовы времени такие, как пандемия COVID-19 и рыночные условия влияют на перевод ряда работ из офлайн-режима в онлайн.

Даже в такой трудно виртуализируемой области, как спорт, клиенты могут участвовать в тренировках онлайн, вместо того чтобы ходить в тренажерный зал [2].

Экономическое преимущество виртуализации бизнеса можно показать на следующем примере.

В «Ростелеком» работают около 150 тыс. человек, треть которых перешла на удаленную работу весной. Производительность труда не снизилась – Ростелеком выполнил даже больший объем работ, чем ранее [4]. Выручка Ростелекома за первое полугодие 2020 года выросла на 10 % по сравнению с аналогичным периодом прошлого года [1, с. 20]. После отмены карантина значительная часть сотрудников продолжит работать удаленно. Другие будут переведены на смешанный график [4].

Цифровые технологии представляют одну из основ бережливого производства [6] (концепции производства, предусматривающей максимальное сохранение ресурсов).

Кроме того, удаленную работу удобно сочетать со фрилансом (*англ.* freelancer – свободный работник).

С экономической точки зрения, преимущества фриланса следующие [8]:

1. Экономия на налогах (нет социальных вычетов);
2. Привлечение работника для решения конкретных задач, а не на постоянную службу;
3. Широкий выбор кандидатов (нет ограничения по городам и странам);
4. Рыночные условия для образования цены. Часто фрилансеры устанавливают небольшую стоимость услуги, а зарабатывают за счет больших объемов и оптимизации своего труда.

Для успешной работы такой команды нужно подобрать грамотный персонал и организовать систему управления проектами. Чтобы сделать это, нужно выполнить ряд условий [9]:

1. Привлекать к оценке проекта членов команды (в свою очередь, их нужно подобрать оптимальным образом);
2. Изучить процессы управления проектами в компании, внедрять современные методологии управления проектами, например, Agile–минимизацию рисков путём сведения разработки к серии коротких циклов, называемых итерациями, которые обычно длятся две-три недели;
3. Обучать команду;
4. Знать проектную историю своей команды;
5. Правильно ставить вопросы и цели;
6. Применять иерархическую структуру работ (здесь пригодятся сервисы проектного управления).

Решить эти задачи можно и нужно с помощью существующих сервисов для обеспечения работы виртуальных команд.

Во-первых, нужно рассмотреть программное обеспечение подбора персонала.

В обзоре ПО найма за 2020 год [10] указаны следующие продукты:

1. Hrscanner – это интернет-сервис для оценки кандидатов и персонала, предназначен для тестирования;
2. Talantix– сервис онлайн-рекрутинга;
3. Робот Вера – предназначенный для подбора персонала цифровой бот (производит поиск кандидатов по сайтам, интервьюирует их, дает оценки);
4. МегаФон Рекрутика – автоматически отбирает с разных сайтов персонал и анализирует;
5. HeadHunter – одна из крупнейших платформ поиска сотрудников и публикации вакансий;
6. VCV – сервис тестирования и интервьюирования с интерактивным чат-ботом;
7. Jobvite – комплексный онлайн-сервис подбора персонала;
8. Huntica – созданный для виртуальных собеседований онлайн-сервис, помогающий работодателю массово интервьюировать кандидатов по подготовленному сценарию при помощи кадрового бота;
9. 1С: Кадровое агентство – позволяет вести учет событий, кандидатов, их ранжирования, контролировать финансовую отчетность;
10. Яндекс.Работа – агрегатор вакансий;
11. HR ЧатБот – цифровой бот для поддержки и обратной связи с сотрудниками;
12. Хантфлоу – распознает резюме, ведет переписку с кандидатами, составляет отчетность по вакансиям;
13. Workable;
14. Hurma System – веб-сервис с возможностью адаптации и сопровождения сотрудников, учета рабочего времени, статистики и HR-аналитики;
15. Cleverstaff – оптимизирует найм;
16. Potok;
17. SuperJob–российский сервис размещения и поиска вакансий, аналогичный Яндекс.Работа, один из наиболее популярных.

Другие сервисы этой категории – это биржа фриланса (место встречи работодателя и работника) Kwork; аналогичные онлайн-сервисы рекрутинга, например, iCIMS Recruit [16] с применением процедур, основанных на искусственном интеллекте, привлекающая значительные инвестиции платформа Turing [12] (преимущество - автоматизированное определение квалификации персонала на основании тестирования, отбор сотрудников, поддержка видеоинтервью) и т.д.

Вторая категория программных продуктов – это системы проектного управления – со своими подкатегориями [7,11,13,14,15,17].

В компаниях используются общепринятые сервисы для управления проектами, например, MicrosoftProject, Covisint, Project Kaiser и др.

Microsoft Project – помогает руководителю проекта в разработке планов, распределении ресурсов по задачам, отслеживании прогресса и анализе объемов работ. Создает расписания критического пути. Поддерживает диаграммы Ганта (столбчатые диаграммы, отображающие планирование проекта – каждый столбец соответствует задаче, его «высота» - времени на решение). Это инструмент, где для внедрения управления проектами надо выбирать методологию проектного управления, реализуемую, как правило, через «регламенты» проектного управления и отраслевые доработки MS Project.

Project Kaiser предлагает аналогичный функционал, но в несколько упрощенном виде (например, без критического пути).

В обзоре [13], выполненном изданием vc.ru, выделяются следующие 5 лучших сервисов управления проектами 2020 года.

1. Asana (позволяет осуществлять мониторинг активности, доступа и приоритетов), поддерживает управление календарем, файлами, отслеживание активности, управление документами, управление разрешениями, контроль доступа пользователей, определение приоритетов задач, управление шаблонами, отслеживание целей, отслеживание этапов и API;
2. ProofHub (предусматривает чаты и обсуждения для пользователей, здесь также применены диаграммы Ганта и доска Канбан-доска);
3. Средство с высокой степенью сотрудничества Basecamp (применимо для предприятий всех размеров, используются шаблоны проектов и совместный доступ к документам). Есть возможность чатов с обсуждениями проектов и документов. Сервис платный, пробная версия – 30 дней;
4. Облачное средство визуального планирования Proggio;
5. Централизованная платформа (онлайн-сервис) Hive. Программа Hive располагает возможностью заполнять сведения о проектах;

Недостаток перечисленных сервисов – отсутствие русскоязычной поддержки (кроме Asana).

Также стоит отметить программу TrackStudio, где задачи собираются на рабочем столе системы и одновременно имеют представление в виде дерева. В TrackStudio отсутствует удобный механизм завершения задач, чек-листов, быстрых моделей управления задачами (назначения, делегирования, контроля и проч.).

Весьма широко используется Jira - коммерческий продукт, система отслеживания ошибок, написанная на Java и имеющая веб-интерфейс. Основной элемент здесь - это задача со своими характеристиками, свойствами и дополнительной информацией. Выполнение задач протоколируется и записывается журнал. Благодаря дополнительным функциям Jira может быть интегрирована с другими средами разработки.

Аналогичный пример: Redmine, с простой процедурой создания задач (вводится название, поле описания, степень готовности, трудозатраты, дедлайн, статус). Менять статус выполнения (в том числе закрывать) можно только посредством редактирования задачи. Чек-листы не предусмотрены. Все задачи отображаются в виде календаря и на диаграмме Ганта.

Другие программы управления задачами – Битрикс, Мегаплан и т.д. Отдельная категория программ – системы, где таск-менеджер – модуль. На них мы останавливаться не будем и перейдём к следующей категории.

Перечислим программы, ориентированные непосредственно на управление удаленными командами и наиболее подходящие для работы с ними.

Socoso (аббревиатура от SocialCommunicationsCompany) – специальное средство найма удаленных сотрудников. Интересна достижением эффекта присутствия за счёт визуализации хорошо знакомого рабочего помещения перед глазами пользователя (наиболее частый пример — офис, в котором он работает) и других сотрудников компании в нём.

Podio - виртуальная рабочая среда с привязкой проектов к клиентам, панелью рабочего пространства с контролем выполнения работы по срокам, чатами для совещаний.

ProWorkflow помогает отслеживать ход и обновлять данные сразу ряда проектов нескольких сотрудников – одновременно и в одном месте. Примеры функций сервиса – отображение рабочей загрузки команды для удобства просмотра в инструментальной панели, а также временная шкала и табель, с которыми учёт затрат времени вести легче лёгкого.

Worketc. Сервис обеспечивает предварительный просмотр условий перевода бизнеса из разряда мелких в средние и в крупные.

Gorplan – система, позволяющая, так же, как и BaseCamp, работать с проектами удаленно. оснащается модулями для коллективного редактирования записок, ведения календарей (экспорт в iCal), управления задачами, отслеживания текущих дел, управления файлами, а также имеет встроенный чат (поддержка SSL).

RingCentral – сервис видеоконференций, позволяющий управлять сотрудниками онлайн.

Преимущество продуктов управления удаленными проектами с разным функционалом состоит в возможности организовать сотрудников и управлять большими объемами информации о проектах, помочь руководителю сориентироваться в принятии решения и организовать производственный процесс оптимальным образом.

В то же время, в большинстве этих систем отсутствует возможность прогнозирования хода проекта и сроков его завершения. В материалах [3, 5] отмечается, что, например, в Microsoft Project встроена поддержка метода PERT, позволяющая определить приблизительный срок выполнения проекта и вероятность выполнить его раньше или позже срока. Но точность такой оценки очень низка в связи со сложностью построения сетевого графика всего проекта и внесения изменений в него в ходе проекта.

Отдельный аспект – это создание среды, обладающей должным уровнем информационной безопасности.

Библиографический список

1. Зима И. Нет ничего важнее доверия клиентов // National Business, август-сентябрь 2020, с. 20 – 22.
2. Как коронавирус изменит работу будущего: 6 трендов // National Business, август-сентябрь 2020, с. 28 – 29.
3. Культин Н.Б. Управление проектами: инструментальные средства. СПб.: Политехника, 2002. 216 с.
4. Осеевский М. События, произошедшие в последние месяцы, кардинально изменили мир // National Business, август-сентябрь 2020, с. 19.
5. Полицын С.А. Система планирования проектов разработки программного обеспечения // Вестник ВГУ, серия: системный анализ и информационные технологии, 2013, № 1
6. Цифровая основа бережливого производства // National Business, август-сентябрь 2020, с. 23.
7. 10 сервисов совместной работы для управления проектами // [Электронный ресурс] Режим доступа: <https://infogra.ru/design/10-servisov-sovmestnoj-raboty-dlya-upravleniya-proektami>(дата обращения: 4.02.2021)
8. Как нанять программиста дистанционно — Майк Лебедев — Хайп // [Электронный ресурс] Режим доступа: <https://hype.tech/@id7711/kak-nanyat-programmista-na-frilanse-l5yghqg>(дата обращения: 4.02.2021)
9. Как точно оценить проект за 6 шагов / Блог компании OTUS. Онлайн-образование / Хабр // [Электронный ресурс] Режим доступа: <https://habr.com/ru/company/otus/blog/524410/>(дата обращения: 4.02.2021)
10. Лучшие Сервисы подбора персонала (RMS) - 2020, список программ // [Электронный ресурс] Режим доступа: <https://soware.ru/categories/recruitment-management-services> (дата обращения: 4.02.2021)
11. Мы обзрели: выбираем систему управления задачами для web-студий / Блог компании PTYSH / Хабр // [Электронный ресурс] Режим доступа: <https://habr.com/ru/company/ptysh/blog/297824/>(дата обращения: 14.09.2020)
12. Платформа для найма удаленных разработчиков Turing привлекла более \$300 тысяч // [Электронный ресурс] Режим доступа: <https://rb.ru/news/turing-angelsdeck/>(дата обращения: 4.02.2021)

13. Пять лучших программ управления проектами на 2020 год — Сервисы на vc.ru [Электронный ресурс] Режим доступа: <https://vc.ru/services/108319-pyat-luchshih-programm-upravleniya-proektami-na-2020-god> (дата обращения: 14.09.2020)
14. Сервисы для онлайн-конференций // [Электронный ресурс] Режим доступа: <https://www.gd.ru/articles/11225-servis-dlya-konferentsiy> (дата обращения: 14.09.2020)
15. Технологии управления удаленными надомными сотрудниками – обзор // [Электронный ресурс] Режим доступа: http://www.livemanagement.ru/tags/organizacija_udalenoj_raboty (дата обращения: 14.09.2020).
16. ТОП-10 онлайн платформ для рекрутеров // [Электронный ресурс] Режим доступа: <https://www.talentscan.pro/ru/blog/platforms/> (дата обращения: 4.02.2021).
17. Online project management software - CitrixPodio // [Электронный ресурс] Режим доступа: <https://podio.com/site/use-cases/project-management> (дата обращения: 4.02.2021)

D.V. Baranov

OVERVIEW OF SOFTWARE PRODUCTS FOR SUPPORTING DISTRIBUTED TEAMS

Nizhny Novgorod state technical university n.a. R.E. Alekseev

The article discusses the work of virtual freelance teams and its algorithm. The features of the required software and its classification are indicated. The directions of development of these software products are outlined.

Keywords: virtual organization, competency assessment, decision support systems, employment testing.

МЕТОДОЛОГИЧЕСКИЕ АСПЕКТЫ ПОДДЕРЖКИ ПРИНИМАЕМЫХ РЕШЕНИЙ ПРИ УПРАВЛЕНИИ ЖИЗНЕННЫМ ЦИКЛОМ ИНФОРМАЦИОННЫХ СИСТЕМ В ФИНТЕХЕ

Нижегородский государственный технический университет им. Р.Е. Алексеева

Рассматриваются организационные аспекты и технологии моделирования процессов управления жизненным циклом информационной системы в финтехе, методологические аспекты управления развитием информационных систем и их жизненного цикла.

Ключевые слова: информационная система, управление развитием информационной системы, жизненный цикл информационной системы.

Современные темпы роста объемов информации, требующей обработки в режиме реального времени, а также тенденция к созданию все большего числа удаленных рабочих мест в различных отраслях экономики заставляют пересмотреть взгляды на организацию процесса управления жизненным циклом информационных систем.

Актуальность создания методов описания и моделирования информационных потоков вызвана высокой степенью интенсификации обмена информацией и распространенностью применения новых информационных технологий. Построение моделей процессов управления жизненным циклом (ЖЦ) информационных систем (ИС) призвано обеспечить поддержку лицам, принимающим решения (ЛПР) для своевременного вмешательства в развитие информационной системы (ИС) компании.

Своевременное выявление тенденций, способных негативно повлиять на ЖЦ ИС финансовой организации, их устранение путем ее доработки и обновления позволит не только адекватно рассчитать необходимые ресурсные затраты на поддержку системы, но также может послужить сигналом к необходимости внесения изменений в устоявшиеся бизнес-процессы организации [4].

Несмотря на то, что в общем виде ИС включает практически все ресурсы организации, так или иначе обрабатывающие информацию или участвующие в информационных потоках, наиболее важное значение имеет та часть ИС финансовой организации, которая обеспечивает автоматизацию основных функций бизнес-подразделений. Ее принято называть автоматизированной банковской системой или корпоративной информационной системой (КИС). Именно КИС рассматривается авторами, как объект исследования, моделирование жизненного цикла которой может быть экстраполировано на ИС финтеха.

Модель жизненного цикла КИС финансовой организации – это комбинация последовательности этапов жизненного цикла и переходов между ними, необходимых для гарантированного достижения поставленной для реализации проекта цели. Специфика ЖЦ КИС финтеха заключается в том, что начальные стадии ЖЦ, такие как обследование предметной области, постановка задачи, проектирование и разработка в настоящее время уже не имеют смысла, т.к. новых финансовых организаций открывается крайне мало, но и для них на рынке существует множество готовых автоматизированных систем, внедрение которых значительно сокращает период запуска КИС. Развертывание и внедрение для финтеха весьма сложные процессы, поскольку занимают значительное время, при этом бизнес должен оставаться непрерывным. Вероятность возникновения необходимости утилизации, текущей КИС и внедрения новой призвано минимизировать для финтеха настоящее исследование в стремлении сохранить текущую КИС. На этапах Эксплуатации, поддержки и модернизации КИС мы и сосредоточимся. Принимаемые на них решения

должны обеспечить максимальное соответствие КИС требованиям бизнеса при минимальных издержках.

Эксплуатация КИС, являющаяся основным этапом, баланс которого необходимо поддерживать, обеспечивается со стороны ИТ-подразделения организации за счет стабильности работы оборудования, системного программного обеспечения и поддержки конечных пользователей. Интеллектуализация принимаемых решений при управлении ЖЦ КИС на данном этапе является особенно важной, поскольку позволяет предотвратить серьезные последствия, связанные со снижением качественных и количественных показателей работы КИС по сравнению с эталонными, соответствующими меняющимся требованиям бизнеса. [1].

1. Постановка задачи поддержки и модернизации корпоративной информационной системы

1.1. Критерии оценки соответствия корпоративной информационной системы требованиям бизнес-заказчика

В своей деятельности сотрудники организации руководствуются определенными регламентами, описывающими бизнес-процессы, которые могут быть оценены с точки зрения затрачиваемого времени, а значит и уровня автоматизации [3]. Последующий мониторинг соответствия данной оценки среднему или эталонному значению может выступать важным показателем потребности внесения изменений как в КИС, так и в сам бизнес-процесс. Многокритериальная оценка и мониторинг показателей основных бизнес-процессов позволит избежать критических ситуаций, когда доработка КИС может оказаться либо экономически не целесообразной или низко эффективной с точки зрения достижения требований бизнеса

Помимо временного показателя рассматриваются степень автоматизации и формализации выполняемых сотрудниками задач. Снижение данного показателя влечет дополнительные затраты по времени на ручную обработку документов.

Удобство интерфейса КИС, а, следовательно, и трудоемкость при выполнении своих бизнес-функций сотрудниками наиболее трудно уловимый фактор, однако, существенно влияющий как на скорость обработки информации, так и на эффективность работы организации в целом. Рутинные действия перестают критически оцениваться исполнителями, но вновь принятые сотрудники могут обратить внимание на низкий уровень автоматизации бизнес-процессов. В то же время, анализ и контроль возникновения подобных ситуаций может быть автоматизирован.

Запуск процессов доработки и адаптации КИС к быстроменяющимся требованиям бизнеса возможен лишь посредством постоянного мониторинга множества факторов и процессов, происходящих в КИС. Однако, для этого необходимо максимально формализовать данные факторы и процессы.

1.2. Основные подходы к оценке релевантности и методы построения КИС

Необходимо выбрать такое концептуальное решение (конфигурацию), которое при максимальной полезности (или качестве) ИС обеспечивает эффективность бизнес-процесса, при этом выполнение требований является необходимым условием, но не гарантирующим достижения максимальной эффективности (соотношение результативности и затрат).

Полезность– степень удовлетворенности пользователей набором функций (модулей) КИС. Поскольку доработка КИС характеризуется заданной трудоемкостью и рисками (проектными или техническими), то эффективность можно оценивать с помощью составного критерия, включающего в себя такие показатели, как U_i – полезность i -го компонента КИС (например, модуля, выполняющего в общем случае одну функцию), T_i – трудоемкость реализации компонента ИС (задается при бизнес-планировании) и P_i – вероятность успешной реализации компонента ИС (оценивается экспертами).

Поиск системно-технических решений, включая оптимизацию показателей U , T , R отдельных компонент позволит максимизировать полезность КИС, что будет способствовать повышению эффективности бизнес-процесса.

Однако, высокие темпы развития информационных технологий приводят к тому, что КИС требует постоянного совершенствования. На рынке информационных технологий возникает множество новых практик, направленных на активное взаимодействие ИТ-специалистов, как разработчиков, так и администраторов с целью повышения темпов их взаимодействия и повышения эффективности выпуска обновлений КИС. Например, DevOps (от англ. development&operations), дополненный гибким подходом к разработке Agile, ориентированных на использование итеративной разработки, динамическое самоорганизующихся рабочих групп, состоящих из специалистов различного профиля.

Подход DevOps формирует непрерывное взаимодействие сотрудников в ИТ-подразделении, без которого польза от моделирования ЖЦ и интеллектуализация принимаемых решений может свестись к нулю. Если, например, администраторы не успевают развертывать в промышленную эксплуатацию поставляемые разработчиком релизы, происходит накопление обновлений, а бизнес-заказчики не получают необходимый функционал.

Кроме того, последнее время стали популярными так называемые микросервисы (МС), на основе которых строится КИС финансовой организации. МС используются в качестве строительных блоков для автоматизации различных частей бизнес-процессов [5]. Микросервисная архитектура имеет свои преимущества в обновлении и масштабируемости КИС, но также содержит недостатки, связанные с управлением растущим количеством микросервисов.

Современные информационные технологии, позволяющие вывести на совершенно новый качественный уровень процесс доработки и обновления КИС с применением адаптивных алгоритмов и использованием баз знаний значительно сокращают издержки и сроки реализации улучшений КИС и минимизируют риски, связанные с квалификацией персонала.

2. Методы поддержки принятия решений

В качестве инструмента создания математических моделей (ММ) может быть использован метод группового учета аргументов (МГУА) [2] как один из наиболее эффективных методов структурно-параметрической идентификации сложных систем по данным наблюдений в условиях неполноты информации. Для того, чтобы повысить информативность исходных данных и точность решения задач идентификации и оптимизации, необходимо выполнить классификацию объектов. Для классификации целесообразно применять методы кластеризации, позволяющие осуществлять разбиение изучаемой совокупности систем на отдельные классы, внутри которых системы считаются однотипными. В качестве алгоритма кластеризации используется метод k -средних или метод взаимного поглощения. Использование этого метода разбиения на классы позволяет получить подмножества, содержащие объекты, однородные с точки зрения значений характеризующих их свойств. Выделяемые подмножества обладают таким свойством потому, что взаимное расстояние (радиус поглощения) между любой парой объектов подмножества меньше, чем расстояние между двумя объектами разных подмножеств.

Затем осуществляется поиск оптимальных альтернатив в своем классе с помощью метода Парето. Выбор из этого множества альтернативных вариантов оптимального возможен с использованием дополнительного критерия. Для этого разрабатывается композиционный метод выбора альтернатив. Он объединяет решение задачи оптимизации целевой функции и сведение частных предпочтений к общим требованиям, что позволяет сделать выбор более объективным и ускорить расчет.

Целевой функцией при построении КИС является критерий эффективности, который будет построен с помощью разрабатываемого метода формирования интегрального критерия эффективности при помощи набора моделей с признаками их значимости, таким образом увеличивая степень определенности при принятии решений.

Лицо, принимающее решения (ЛПР), в ходе своей деятельности исходит из опыта, интуиции и знаний. Однако, данные основания трудно поддаются алгоритмизации, вследствие чего требуется активная алгоритмическая поддержка выбора при принятии решений путем построения адаптивных моделей.

Важная роль в процессах принятия решений принадлежит экспертам. В то же время компетентность экспертов, как правило, неизвестна, а экспертные оценки могут быть смещены. Эффективными являются технологии поддержки принятия решений, реализованные с применением технологий классификации и оптимизации вариантов, а также с применением нечеткой логики.

Заключение

Требования к КИС, как и к ИС в целом в финансовой организации постоянно меняются, что накладывает дополнительные требования к ее гибкости, безопасности и скорости адаптации. Постоянные изменения законодательства, вносимые регулятором, связаны с риском, который необходимо учитывать в деятельности организации. Эти изменения могут оказать серьезное влияние на бизнес-процессы, а, следовательно, и на саму компанию.

Таким образом, концептуальная модель управления ЖЦ КИС, как наиболее подверженного изменениям элемента ИС финансовой организации, является основой для создания универсальной технологии построения КИС. Данные механизмы позволят отслеживать не заметные на вид негативные тенденции деградации КИС, в том числе в рамках ее интеграции с другими системами, как в монолитной, так и в микросервисной архитектуре. Конечной целью реализации данных моделей является возможность автоматизировать мониторинг соответствия КИС изменяющимся потребностям, в т.ч. в перспективе, и информировать ЛПР о необходимости вмешательства в ЖЦ ИС. Непрерывное совершенствование ИС обеспечит ее гармоничное эволюционное развитие в рамках бизнес-функций финтеха, а не в отрыве от них.

Библиографический список

1. Дмитриева Н.Г. Методологические модели управления информационным обеспечением // Труды НГТУ им. Р.Е. Алексеева. 2017. № 1 (116). С. 11-22.
2. Ивахненко А. Г. Моделирование сложных систем по экспериментальным данным / А.Г. Ивахненко, Ю.П. Юрачковский. – М.: Радио и связь, 1987. – 120 с.
3. Квасов А.С., Методологические аспекты обеспечения соответствия информационной системы требованиям бизнес-заказчика. ИСТ-2017. – Нижний Новгород, 2017. – С. 51-54.
4. Квасов А.С. Анализ процессов управления жизненным циклом информационных систем / А.С.Квасов // Материалы XXV Международной научно-технической конференции ИСТ-2019. - Н.Новгород: НГТУ им. Р.Е. Алексеева, 2019. – С. 298 - 301.
5. Квасов А.С. Многокритериальный анализ соответствия информационной системы требованиям бизнес-заказчика Материалы XVII Международной молодежной научно-технической конференции “Будущее технической науки”. — Н. Новгород: НГТУ им. Р.Е. Алексеева, 2018. – С. 35-36.

6. Методы проектирования информационно-управляющих и телекоммуникационных систем / Под ред. В.Р. Милова, В.Г. Баранова. – М.: Радиотехника, 2016. – 216 с.

A.S. Kvasov

**METHODOLOGICAL ASPECTS OF SUPPORTING THE DECISIONS MADE IN
MANAGING THE LIFE CYCLE OF INFORMATION SYSTEMS IN FINTECH**

Nizhny Novgorod State Technical University n.a. R.E. Alekseev

The organizational aspects and technologies for modeling the processes of managing the life cycle of an information system in fintech, methodological aspects of managing the development of information systems and their life cycle are considered.

Keywords: information system, information system development management, information system life cycle.

НЕКОТОРЫЕ СРАВНИТЕЛЬНЫЕ ЭКСПЕРИМЕНТЫ ДЛЯ ЭЛИНГИ

Нижегородский государственный технический университет им. Р. Е. Алексеева

Рассматривается сравнение эффективности поиска решения по тексту на основе стандартных средств ИТ-технологий и автоматизированным способом при помощи нового варианта интеллектуальной электронной книги - элинга. Применяется база знаний на основе молинг для текста, используемого в эксперименте. Исследование демонстрирует более высокое качество нахождения решения и потенциально более высокую скорость поиска решений.

Ключевые слова: эксперимент, элинга, база знаний, текст, знания, диалоговая система, молинга.

Введение

При принятии решений в сложной ситуации человек должен учитывать структуру связей и зависимостей, которые еще меняются во времени. Эффективным орудием познания таких ситуаций и систем сложных систем, характеризующихся комплексностью, динамичностью, неопределенностью, непредсказуемостью, являются не только знания, зафиксированные в виде теоретических понятий разной степени абстрактности, строгости и точности, но и понятия нестрогие и нечеткие, построенные на основе эмпирических, а не теоретических, обобщений, а также динамические образные представления, которые сложно фиксировать в виде строгих и точных понятий и устойчивых классификаций [1 - 4]. При исследовании сложной системы необходимо множественное целеполагание — постановка разнообразных, разнотипных и разноуровневых целей, связанных с различными подсистемами, аспектами изучаемой комплексной динамической ситуации. Постановка одной цели принципиально недостаточно, сколь бы конкретной или, наоборот, общей она ни была. Часть этих разнообразных целей неизбежно конкурирует между собой (как минимум, за отводимое на их достижение время).

Для облегчения принятия решений человеком в процессе мыслительной деятельности предназначены ряд направлений развития интеллектуальных систем (ИС) с разной степенью успешности и полноты - экспертные системы (ЭС) [5], ИС ALICE Ж.-Л. Лорьера [6], ИС на основе eSAO-структур [7], а также ИС на основе технологии прямого наложения знаний (ТПНЗ) [8 - 10].

Важной задачей для определения, что лучше подходит для решения подобных проблем можно использовать аналитические методы, изучение практики массового применения и экспериментальные методы. Практику массового применения сложно использовать для вновь создаваемых ИС, таких как ИС на основе ТПНЗ. Некоторые результаты аналитического сравнения уже приводились в [11], но наиболее показательны конечно экспериментальные методы [1 - 4].

В результате была проведена серия экспериментов на основе когнитивного подхода с использованием числового шкалирования по сравнению качества работы человека при работе с научно-техническими текстами с использованием традиционных современных методов работы на основе, скажем, пакета Microsoft Office и средств Интернета с возможностями работы с теми же текстами, но уже промоделированными и превращенными в базу знаний (БЗ) из молинг, с применением прототипа ИС на основе ТПНЗ - элинги [8-10].

Описание и методика проведения эксперимента

В рамках данного исследования проводилась эффективность использования элинги при решении ряда поставленных задач, при этом пользователю требовалось инициировать диалог с элингой, сформировать запросы, которые позволят решить поставленные задачи,

путем получения ответов из БЗ посредством *элинги*. Элинги представляет собой ИС, использующие БЗ из молинг, полученные на основе ТПНЗ [8 - 10]. Они предназначены, в основном, для применения индивидуальным потребителем. В состав элинги входят программные средства, реализующие машину вывода, интеллектуальный интерфейс, БЗ, сервисные и вспомогательные программы. БЗ включает в себя словари терминов, отношений и др. (и их синонимов), собственно знания, содержащиеся в текстах в виде молинг, расчетные модели (реализованные в виде отдельных подпрограмм, ускоряющих работу элинги, и повышающих ее ценность), визуальные и графические образы. Структурная схема элинги показана на рис. 1.

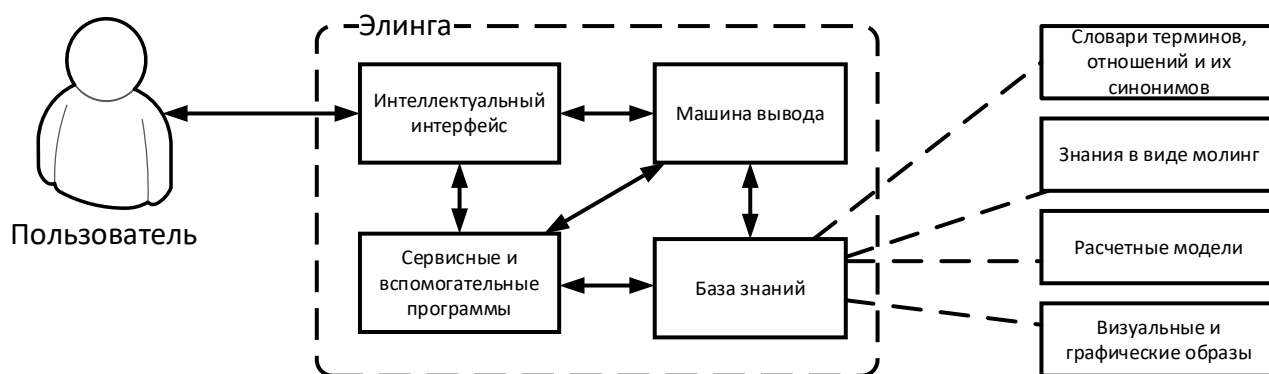


Рис. 1. Структурная схема элинги

БЗ представляет собой знания, относящиеся к некоторой предметной области и формально представленных таким образом, чтобы на их основе можно было осуществлять мыслительный процесс. БЗ впервые начали использоваться в контексте создания ЭС, где в них аккумулируются навыки и опыт экспертов, занятых практической деятельностью в соответствующей области (например, в медицине или в математике). В ЭС БЗ обычно содержат описания в форме конкретных фактов и правил логического вывода над данными и процедурами обработки информации, представляющих сведения и знания о людях, предметах, фактах событиях и процессах в логической форме.

Для описания знаний в элинге используется модель знаний, названная *молингой*, и заключается в следующем. Каждое простое предложение текста (сложные делятся на простые) представляется маленькой семантической сетью с составлением словарей терминов (включая синонимы), отношений, качественных и количественных признаков и т. д. Из предложений убирается эмоциональная окраска, они упрощаются с сохранением основных риторических отношений [8 - 10].

Элинги позволяют: заносить в них постепенно большое количество знаний; пользователю получать ответ сразу (почти со скоростью расчетных программных средств) без изучения технической литературы, задавая вопросы элинге; работать в широкой предметной области; пользователю постепенно получать со временем результаты на высоком уровне. В процессе взаимодействия пользователь может сохранять результаты логического вывода программы и на основании этих данных формировать отчеты (рис. 2).

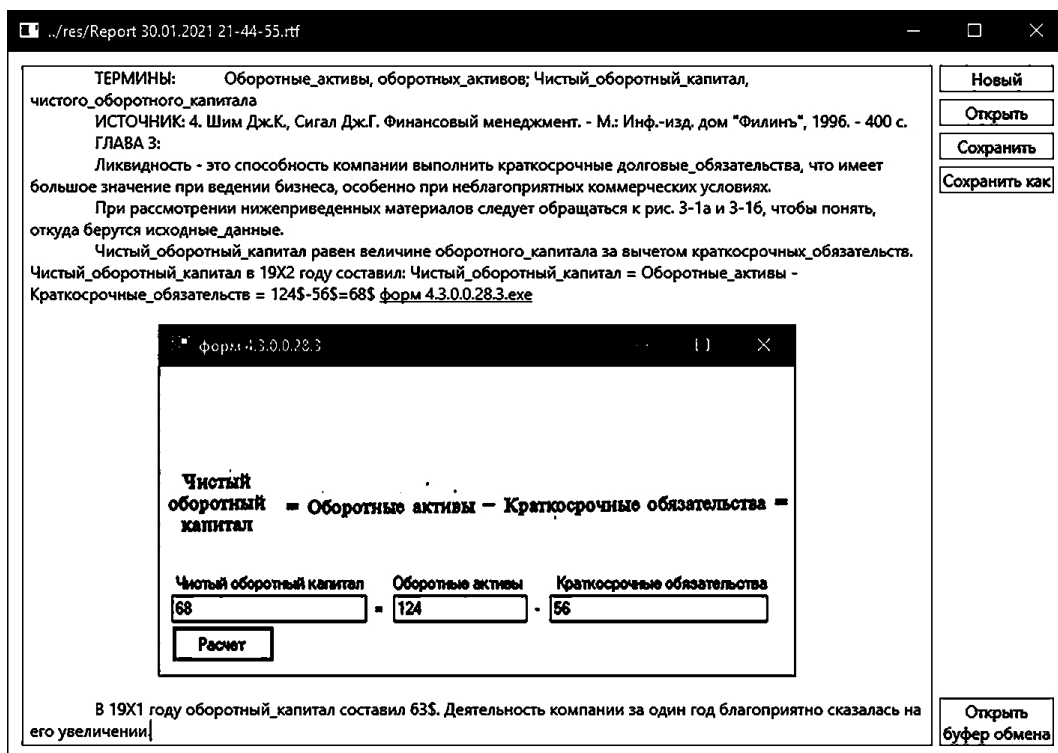


Рис. 2. Отчет, полученный с помощью элингой в результате диалогово-ассоциативного поиска

Последние несколько лет мы упорно делали БЗ из молинг силами студентов в рамках учебного процесса на основе ТПНЗ. Постепенно разработали полуавтоматический редактор знаний (РЗ) для моделирования текстов [12], он немного ускорил и облегчил работу, но принципиально полно решить проблему не может, как это не удалось пока никому.

Сосредоточились пока на двух книгах из области экономики и менеджмента. Одну из них почти промоделировали до конца [13]. Она хороша тем, что была у меня в наличии, ясно и четко изложен исходный материал, много расчетных формул достаточно простых, но полезных, и почти все с числовыми примерами (что вообще редкость). Вторая книга – авторы посолидней намного, но формул очень мало и гораздо объемнее. Поскольку элинга позволяет запускать расчетные формулы и демонстрировать рисунки, то создание БЗ на основе одного источника – означает, что мы сделали новый «электронный учебник», почти аналогичный [14]. Может быть в несколько необычной форме, но именно так, каждый раз новый для какой-то новой области в зависимости от предметной направленности БЗ.

Была неясность, что с чем сравнивать. Литература по результатам сравнения как решает человек задачи без ЭВМ и с применением ЭВМ есть, например, [1]. Но есть давно и другие информационные системы и ИС. Об этом у нас не раз было сравнение с элингой, в частности, [10-11]. Но это в теории. Решили сделать экспериментально на практике, причем достаточно давно, но необходимо совпадение сразу нескольких условий.

Проблему вначале создавало дистанционное обучение (ДО), но потом, пожалуй, оно только пользу принесло для этой ситуации, поскольку отсутствовал практически полностью контроль за поведением студентов в этой ситуации.

Вначале выдали студентам полный скан книги [13] не менее чем за неделю, попутно выяснилось, что книга и в Интернете стала доступна. Предупредили, что будет 2 вопроса, первый из конкретных названных глав, рекомендовалось их изучить подробнее. 2-й вопрос из других глав (всего их 18). Все вопросы с обязательным применением расчетов. 2-й задавался, когда ответят на первый.

Книга по экономике, студенты – это магистры по информационным технологиям, предмет для них не профильный, но что-то по экономике у них было. В книге материал более сложный и конкретный, в целом – они его не знают. Но студенты люди привычные к быстрой подготовке и поиску в текстах многими годами учебы. ДО исключало контроль за ними, поэтому они могли применять любые доступные им методы поиска и решения, в т.ч. и в подготовке для этого. Будем называть это автоматизированным поиском (АП-поиск) – сочетанием ручных способов и доступных студентам программных средств (и их возможностей).

Замеряли время ответа и качество решения (в сравнении с источником) для каждого, кто участвовал в ответах. Участвовало около 19 человек из 4 групп (группы малочисленные).

Через 2 недели повторили то же при ответах с помощью прототипа элинги. Но тут сложнее – мы все равно проходили изучение прототипа элинги в конце семестра и у них у всех малый опыт работы с ней. Это создавало и более лучшие и более худшие условия для сравнения. А возможно это почти идеальная ситуация для сравнения.

Результаты проведения исследования показаны в табл. 1-2. Так, в табл. 1 показаны средние оценки по группам пользователей, табл. 2 отображает среднее время ответа на вопросы по группам.

Таблица 1

Оценка и время при АП-поиске ответов

	Оценка из 11 баллов		Время, мин		
	Вопрос 1	Вопрос 2	Вопрос 1	Вопрос 2	Общее
ГРУППА1 (4 студента)	7.60	6.40	7.00	11.80	18.80
ГРУППА2 (6 студентов)	6.00	6.50	16.50	17.00	33.50
ГРУППА3 (2 студента)	6.75	7.00	18.00	12.75	30.75
ГРУППА4 (5 студентов)	5.00	6.17	10.17	15.67	25.83

Таблица 2

Оценка и время при поиске ответов с помощью элинги

	Оценка из 11 баллов		Время, мин		
	Вопрос 1	Вопрос 2	Вопрос 1	Вопрос 2	Общее
ГРУППА1 (3 студента)	9.67	9.67	11.00	12.00	23.00
ГРУППА2 (6 студентов)	7.00	7.00	20.00	15.00	35.00
ГРУППА3 (2 студента)	9.67	9.67	14.33	9.00	23.33
ГРУППА4 (6 студентов)	8.83	9.33	10.17	13.00	23.17

При анализе данных таблиц можно увидеть, что средняя оценка ответов на вопросы при использовании элинги по всем исследуемым группам выше, чем при АП-поиске ответов и находится в интервале от 7,00 – 9,67 баллов, тогда как при АП-поиске ответов интервал оценок на вопросы от 5,00 до 7,60 баллов. Это указывает на то, что с помощью элинги пользователи находили более точные ответы на поставленные вопросы. По поводу средней скорости поиска ответов (на два вопроса) с помощью элинги по среднему показателю всех групп, результаты, полученные с помощью элинги заняли меньше времени – 12,26 мин (со средне квадратичным отклонением – 6,93 мин), при АП-поиске получение результата заняло – 12,91 мин (со средне квадратичным отклонением – 6,69 мин).

Это первый пробный этап экспериментов для отработки методики. В дальнейшем продолжим.

Пока лишь краткие выводы

1. Главный – элинга работает даже в режиме исследовательского прототипа и выполняет свои функции, как, собственно, и было впервые заявлено 19 лет назад в 2002 г. в Интернете и далее [8-11].
2. Среднее время ответов приблизительно одинаковы – 12,91 мин – для АП-поиске ответа по тексту (разброс от времени ответов от 5 мин до 30 мин), для поиска решения с помощью элинги 12,26 мин (разброс от времени ответов от 3 мин до 30 мин). Но тут надо иметь в виду, что некоторые студенты действительно почти впервые использовали элингу и иногда отвлеклись на домашние дела, поскольку эксперимент проводился в рамках ДО.
3. Качество ответов через элингу было намного выше, чем при АП-поиске ответов. Некоторые группы почти все ответили на отлично. Удивительно, что не было ни одного серьезно ошибочного ответа при применении элинги. Тогда как при АП-поиске ответов несколько студентов серьезно ошиблись с ответами.
4. На удивление, не было ни одного отказа программного обеспечения прототипа элинги при работе с неподготовленным пользователем. Было более 40 попыток решения (не менее), некоторым давались новые вопросы и не один раз из-за проблем с БЗ. К сожалению, БЗ накапливалась долгими годами, в ее подготовке участвовали разные студенты с разным отношением к учебе и в отдельных местах периодически встречаются неожиданности. Но это не приводит к зависанию системы. И потом мы просто постепенно уточняем описание конкретных молинг в БЗ. У нас при работе используется свой РЗ [12], а он всегда надежно проверяет структуру БЗ, поскольку это может привести к проблемам с функционированием самой элинги.

В разработке методики и проведение экспериментов принимал активное участие магистрант Е. Н. Самарин.

Заключение

При проведении исследования была разработана методика, которая позволяет определять эффективность интеллектуальных электронных книг в форме элинга, ускоряющих решение определенных задач при серьезном повышении качества ответа. Подводя итог проведенного исследования можно сделать вывод, что использование элинги при решении задач является значительно более эффективным решением в сравнении с поиском ответов на поставленные вопросы при использовании ряда современных ИТ-технологий.

Исследование выполнено при финансовой поддержке РФФИ в рамках научного проекта № 20-37-90127.

Библиографический список

1. Тихомиров О.К. Психология мышления. – М.: Издат. центр «Академия», 2007. – 288 с.
2. Когнитивная психология. Учебник для вузов/Под ред. В. Н. Дружинина, Д. В. Ушакова. – М.: ПЕР СЭ, 2002. – 480 с.
3. Экспериментальная психология /под общ.ред. Романовой Е.С. – СПб.: Питер, 2007. – 496 с.
4. Микони С.В. Многокритериальный выбор на конечном множестве альтернатив: учебное пособие. – СПб.: Издательство «Лань», 2009. – 272 с.
5. Гаврилова Т.А. Инженерия знаний. Модели и методы / Т. А. Гаврилова, Д. В. Кудрявцев, Д. И. Муромцев. – СПб.: Изд-во «Лань», 2016. – 324 с.
6. Лорьер Ж.-Л. Системы искусственного интеллекта / Ж.-Л. Лорьер. – М.: Мир, 1991. – 568 с.
7. IHS Goldfire. Accelerating Decisions. Powering Innovation. White Paper// IHS Inc., 2013. 7 p.

8. Бронфельд Г.Б. Система для работы с интеллектуальной электронной книгой – элингой – носителем знаний//Патент на полезную модель, № 104348, МПК G06 N 5/00, приоритет от 06.07.2010 г. опубликовано 10.05.2011. Бюл. №13
9. Бронфельд Г. Б. Основы искусственного интеллекта. – Н. Новгород: НГТУ им. Р. Е.Алексеева, 2014. – 253 с.
10. Бронфельд Г.Б. Прототип интеллектуальной электронной книги с использованием технологии прямого наложения знаний/ Г.Б. Бронфельд, Д.И. Киров, В.В. Кондратьев //Программные продукты и системы, №3, 2019. С.403-410.
11. Бронфельд Г.Б. Сравнение особенностей применения элинги и других интеллектуальных систем // Материалы "Международной научно-технической конференции
12. "Информационные системы и технологии ИСТ-2014", г.Н.Новгород, НГТУ им. Р.Е.Алексеева, 2014. С. 204-205.
13. Киров Д.И., Бронфельд Г.Б. Редактор знаний для интеллектуальной системы с базой знаний на основе молинга// Системы управления и информационные технологии, №2(80), 2020. С. 71-76.
14. Шим Дж.К., Сигал Дж.Г. Финансовый менеджмент. – М.: Инф. -изд. дом "Филинь", 1996. – 400 с.
15. Башмаков, А.И. Разработка компьютерных учебников и обучающих систем /
16. Башмаков А.И. и Башмаков И.А. – М.: Информационно-издательский дом «Филинь», 2003. – 616 с.

D. I. Kirov, G. B. Bronfeld

SOME COMPARATIVE EXPERIMENTS WITH ELINGA

Nizhny Novgorod State Technical University n.a. R.e. Alekseev

The article deals with the comparison of the efficiency of searching for a solution in text based on standard means of IT technologies and based on an automated way using a new version of an intelligent electronic book - elinga. The text used in the experiment is modeled as a knowledge base of molingas. The study demonstrates higher quality of solutions found and potentially higher speed of search.

Keywords: experiment, elinga, knowledge base, text, knowledge, dialogue system, molinga.

АНАЛИТИЧЕСКАЯ СРАВНИТЕЛЬНАЯ ОЦЕНКА СОВРЕМЕННЫХ ВИДОВ БЕСПРОВОДНЫХ САМООРГАНИЗУЮЩИХСЯ СЕТЕЙ AD-НОС

Нижегородский государственный технический университет им. Р.Е. Алексеева

Проводится сравнительно обобщенная оценка вариантов самоорганизующихся беспроводных сетей связи и передачи данных Ad-Нос, представляющих собой децентрализованные беспроводные динамические (не имеющие постоянной структуры) сети. Соединение клиентских (абонентских) устройств осуществляется из условий возможностей (из-за ограниченного расстояния) радиосвязи между этими устройствами или через узлы связи, тем самым создается самоорганизующаяся сеть. На основании этой оценки обосновывается классификация таких сетей в сравнении с другими видами подвижной связи.

Ключевые слова: Ad-Нос сеть, SON, MANET, WSN, WMN, VANET, FANET, OSI, DSRC, V2V, V2I, БПЛА.

Введение

Беспроводная Ad Нос - сеть (от лат. ad hoc - специально, применимо только для этого, т.е. - для специальной цели) - это динамическая, самоорганизующаяся и децентрализованная беспроводная сеть, не имеющая постоянной структуры. Клиентские устройства соединяются в зависимости от условий радиосвязи, образуя собой сеть. Каждый узел сети пытается переслать данные, предназначенные другим узлам. При этом определение того, какому узлу пересылать данные, производится динамически, на основании связности сети. Это является отличием от проводных сетей и управляемых беспроводных сетей, в которых задачу управления потоками данных выполняют маршрутизаторы (в проводных сетях) или точки доступа (в управляемых беспроводных сетях).

Историческим периодом начала создания беспроводных самоорганизующихся сетей можно считать начало 1970-х годов в сети «packet radio» [1].

Указанные свойства самоорганизующихся сетей, благодаря которым реализуется минимальное конфигурирование и быстрое развёртывание, позволяют их применять в условиях экстренных ситуаций и в том числе оперативного мониторинга таких ситуаций.

Целью данной работы является обобщенное аналитическое определение принципов функционирования таких сетей и классифицирование по видам их реализации с оценкой особенностей работы.

1. Основные принципы действия беспроводных самоорганизующихся сетей

Основные принципы действия беспроводных самоорганизующихся сетей можно представить следующим образом. Сети — Ad Нос — это сети, в которых каждый узел может пересылать данные для других узлов, когда определение того, какие узлы пересылают данные, производится динамически на основе возможности сетевого подключения. *Это отличает их от проводных сетей*, в которых, как упоминалось выше, маршрутизаторы выполняют задачу маршрутизации на основе управления в каждой созданной такой структурированной сети. Такие управляемые или структурированные сети, в которых используется специальный узел, известный как точка доступа, управляет связью между другими узлами.

В отличие от таких структурированных сетей, сети Ad Нос формируются на основе того, что узлы сами по себе являются устройствами связи, не нуждаясь в какой-либо инфраструктуре или централизованном управлении. Устройства взаимодействуют

напрямую друг с другом, образуя цепочки передатчиков, тем самым осуществляя передачу информации через другие устройства, чтобы достичь конечного пункта назначения. Устройства также узнают о своих коллегах, а затем используют этот интеллект для маршрутизации информации по оптимальному пути с учетом таких факторов, как вычислительная мощность, емкость аккумулятора и альтернативные сетевые подключения.

Ad-Нос сеть может определять и регистрировать в сети новых подключенных устройств, т.е. осуществляется «самоконфигурация». При этом соседние устройства автоматически корректируют свои технические параметры, такие как мощность излучения, наклон антенны и т.д. Кроме этого адаптируются параметры устройств при изменении параметров сети, т.е. реализуется «самооптимизация» в зависимости от количества пользователей, уровня сигнала, уровня внешних помех. Также осуществляется автоматическое обнаружение и устранение сбоев («самовосстановление») – путем перераспределения функций между устройствами при выходе из строя каких-либо узлов сети, благодаря чему повышается отказоустойчивость сети.

На рис. 1 показана схема мобильных Ad Нос сети с примером подключения новых устройств. Узлы 1,2,3...12 существующие мобильные узлы в сети MANET, узел 13 это новое подключенное устройство.

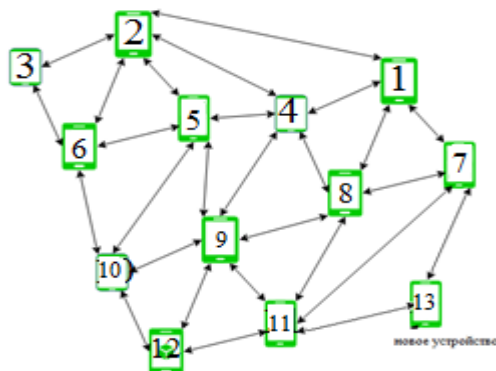


Рис. 1. Мобильные Ad Нос сети (MANET- Mobile Ad hoc Network)

Например, мобильный узел 1 может напрямую связываться с мобильным узлом 2 [однокачковая связь (single-hop)], при условии, что канал доступен. Если канал напрямую между ними недоступен, то необходима многошаговая связь (multihop), например $1 \rightarrow 4 \rightarrow 2$. Для работы многошаговой связи промежуточные узлы должны маршрутизировать пакет данных, то есть они должны действовать как маршрутизаторы. В предыдущем примере узел 4 действует как маршрутизатор. Если требуется продолжение связи - до узла 12, то, поскольку узел 2 не может напрямую связываться с узлом 12 из-за ограниченного расстояния радиосвязи, то для гарантированной доставки пакета данных к месту назначения необходима многошаговая связь, например $2 \rightarrow 6 \rightarrow 10 \rightarrow 12$. В этом случае промежуточные узлы 6 и 10 должны действовать как маршрутизаторы.

При подключении нового узла (устройства) он (на рис. 1 это узел 13) должен быть определен по местоположению и зарегистрирован в сети MANET.

На основании выше изложенного следует, что одним из подходов классификации беспроводных сетей связи является разделение на централизованные инфраструктуры и самоорганизующиеся. Отличительной особенностью самоорганизующихся сетей SON (self-organization) — это возможность в отсутствии централизованной инфраструктуры обмениваться данными любой паре находящихся в зоне радиопокрытия узлов сети. Узлы в SON могут быть одновременно конечными хостами и маршрутизаторами. Соединение организуется на длинные расстояния с помощью специализированных протоколов маршрутизации в промежуточных узлах - маршрутизаторах. Такое соединение называется

«многоэтапным или многошаговым» (multihop). Этапом считается участие в этом соединении одного узла – маршрутизатора [2].

Наглядное различие между централизованно управляемыми сетями и самоорганизующимися сетями мобильной связи демонстрируется на рис.2. Традиционно в многопользовательских системах участники объединяются в ячейку централизованного управления. Базовая станция координирует и предоставляет пользователям доступ к сети, а также распределяет ресурсы (частоты, время и т. д.) между пользователями. Так работают мобильные 2G, 3G, 4G (LTE), планируемые 5G и WiMax сети (рис. 2, а). Такие сети также могут называться *инфраструктурными сетями*.

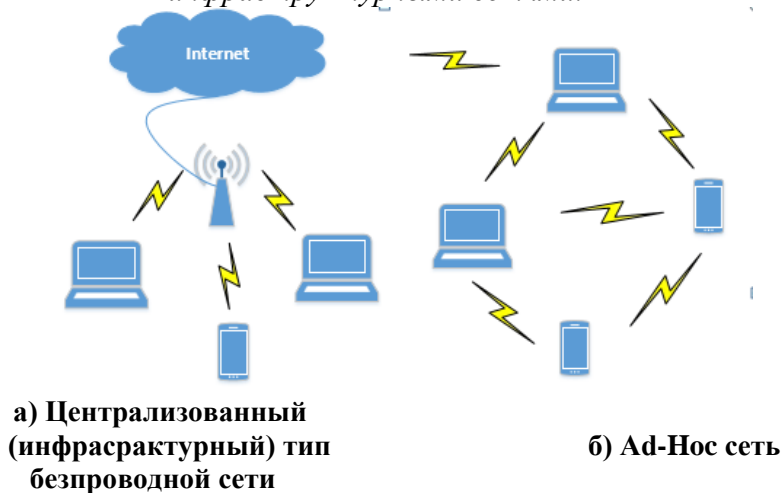


Рис. 2. Разница между централизованной и децентрализованной Ad-Нос сетями

Беспроводные локальные сети (WLAN) также частично следуют этой модели, например, для доступа к Интернету через точку доступа, через которую осуществляется связь между участниками сети (рис. 2, б). Но основной принцип Ad-Нос сети базируется на обмене данными между устройствами (D2D или P2P), при котором устройства обмениваются данными напрямую (рис. 2, б). Обычно Ad-Нос сеть, как указывалось выше, состоит из множества беспроводных устройств, которые обмениваются данными напрямую или в несколько этапов. В отличие от сетей инфраструктурного типа, в ней нет централизованного управления, и все участники сети равны (децентрализованы). Такая сеть должна иметь возможность самоорганизации, которая включает автоматическую настройку программного обеспечения с минимальным участием пользователя. Ad-Нос сеть может быть подключена к другим сетям (включая Интернет) через шлюз. Такую функцию может выполнять, например, устройство, оснащенное модемом для подключения к мобильной сети.

Самоорганизующиеся сети обладают следующими преимуществами над беспроводными сетями традиционной архитектуры: возможность передачи данных на большие расстояния без увеличения мощности передатчика, устойчивость к изменениям в инфраструктуре сети, возможность быстрой реконфигурации в условиях неблагоприятной помеховой обстановки, простота и высокая скорость развертывания

2. Классификация самоорганизующихся беспроводных сетей

Технология беспроводных самоорганизующихся сетей определяется в зависимости от критериев их выбора, которые могут быть классифицированы следующим образом: по иерархии (одноранговые - одношаговые и mesh-сети — сети с ячеистой топологией), по применению (беспроводная сенсорная сеть, транспортная ad hoc-сеть), по мобильности (мобильные самоорганизующиеся сети). Проанализируем функциональные особенности классифицированных таким образом сетей.

2.1. Мобильные Ad Hoc сети - Wireless Mobile Ad Hoc Network (MANET)

Мобильные Ad Hoc сети -MANET состоят из мобильных устройств, обменивающихся данными по беспроводным каналам без какой-либо поддержки со стороны фиксированной инфраструктуры. Такие сети применимы к широкому спектру приложений, которые могут включать в себя аварийное восстановление или создание тактической связи, подключение нескольких мобильных пользователей, использующих ноутбуки, портативные устройства, мобильные телефоны, медиаплееры и т. д. В этой сети пользователи могут обмениваться музыкальными файлами, фотографиями, и т. п. в непосредственной близости, без подключения к стационарной сети или Интернету [3].

Важным свойством MANET является возможность многошаговой передачи (multihop). В отличие от сотовых сетей, которые являются одношаговыми беспроводными сетями (single-hop). Мобильные Ad Hoc сети не гарантируют, что мобильный узел может напрямую связываться со своими пунктами назначения все время. Мобильный узел, который лежит за пределами диапазона передачи своего конкретного пункта назначения, может ретранслировать свой информационный поток через другие мобильные узлы. Это означает, что мобильные узлы в Ad Hoc сетях имеют функциональную возможность маршрутизации и могут действовать как маршрутизаторы и как хосты.

Мобильные Ad Hoc сети -MANET могут быть развернуты в местах, где поддержка мобильной связи недоступна, например, из-за высоких затрат на ее развертывание или аварийных разрушений других сетей связи [1, 4].

2.2. Беспроводные сенсорные сети - Wireless Sensor Network (WSN)

Беспроводные сенсорные сети - WSN могут быть классифицированы как самонастраиваемые и не требующие заданной инфраструктуры беспроводные сети. Применение таких сетей возможно для мониторинга физических или окружающих условий (например, сбора данных о параметрах внешней среды, таких как температура, звук, вибрация, давление, движение или загрязняющие вещества), и для совместной передачи (от узла к узлу) полученных данных через сеть в базовую станцию, где можно наблюдать и анализировать данные. Приемник или базовая станция действует как интерфейс между пользователями и сетью.

Сенсорные сети строятся на основе протоколов IEEE 802.15.4, ZigBee и DigiMesh. С помощью радиосвязи, осуществляемой между узлами сети на основе стандарта ZigBee, создаются самоорганизующиеся и самовосстанавливающиеся сети. Для многих сенсорных сетей характерна мобильность не отдельно каждого узла (как это имеет место в MANET), а отдельной группы узлов. Основное требование к протоколам сенсорных сетей малое потребление энергоресурсов. В сенсорных сетях время их жизнедеятельности прямо зависит от решения вопросов энергопотребления узлов сети.

Сенсорные сети применяются в различных областях. Их можно разделить на категории [5]:

- погода, окружающая среда, сельское хозяйство.
- телемедицина.
- чрезвычайные ситуации (пожары, катастрофы и др.).
- военные операции и др.

2.3. Беспроводные ячеистые сети - Wireless Mesh Network (WMN)

Беспроводные ячеистые сети WMN - это особый вид беспроводных сетей, которые обеспечивают их полную доступность, применяя механизм маршрутизации на каждом узле, который пересылает данные в режиме нескольких переходов. Сетевые маршрутизаторы фиксируются в статичном положении или, по крайней мере, имеют ограниченную мобильность, поэтому создается квазистатическая беспроводная инфраструктура сети. По сравнению с другими беспроводными сетями, такими как мобильные (MANET) или сенсорные сети, узлы WMN не имеют энергетические или

вычислительные ограничения. Кроме того, узлы WMN обычно обеспечивают гораздо лучшее качество связи, чем другие виды сетей, поскольку они могут одновременно использовать несколько радио интерфейсов, разные каналы и более высокую мощность передачи.

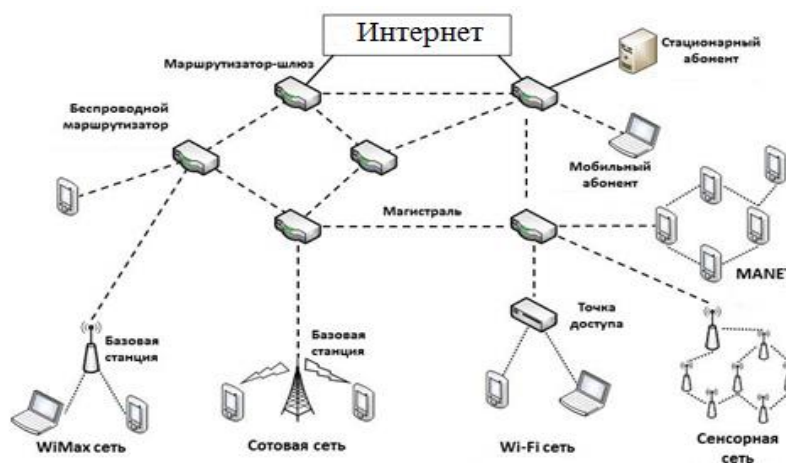


Рис. 3. Архитектура гибридной – ячеистой сети

Такие сети могут взаимодействовать с другими видами беспроводных сетей связи. На рис. 3 приведена общая архитектура гибридной ячеистой сети. Как видно из рисунка, такая сеть состоит из ячеистых маршрутизаторов, которые связываются между собой напрямую друг с другом образуют беспроводную магистраль (Wireless Mesh Backbone) и подключенных к ней сети Интернет, сети Wi-Fi, сотовых сетей связи, конечных пользователей [6].

Следует отметить, что абоненты гибридной ячеистой сети могут получать доступ к сети не только через ячеистые маршрутизаторы, а также напрямую взаимодействовать с другими ячеистыми клиентами (D2D - технология device-to-device позволяет устройствам, находящимся неподалеку друг от друга, обмениваться данными напрямую, без участия сети) [6].

По D2D технологии следует заметить, что она запланирована во внедряемых сетях сотовой связи 5G.

2.4. Автомобильные беспроводные сети -Vehicular Ad Hoc Network (VANET);

Автомобильные беспроводные сети VANET - это тип сетей, созданный на основе концепции создания сети автомобилей для конкретной потребности или ситуации. Основные цели VANET — помочь группе транспортных средств создать и поддерживать сеть связи между ними без использования центральной базовой станции, помочь водителю (это навигация, предотвращение столкновений и смена полос), информирование об ограничении скорости или зоне ремонтных работ, предупреждение о послеаварийных ситуациях, о препятствиях или состоянии дорог и т.п.

Автомобильные беспроводные сети отвечают за связь между движущимися транспортными средствами в определенной среде (зонах). Транспортное средство может напрямую связываться с другим транспортным средством, что называется связью между транспортными средствами (V2V) или транспортное средство может связываться с инфраструктурой, такой как придорожный блок, именуемое связью транспортное средство-инфраструктура (V2I). Сопоставляя D2D в 5G с V2V, следует заметить, что в V2V оба типа узлов представляют собой радиосвязь ближнего действия (DSRC). DSRC может работать в диапазоне 5,9 ГГц с полосой пропускания 75 МГц и с приблизительным интервалом расстояний до 300 м [7]. В основе V2V лежит технология D2D-коммуникаций или иными словами взаимодействия устройство-устройство (Device to Device), рассчитанные на радиосвязь на близком расстоянии и применяемыми для транспортных средств. На рис 4 показан типичный сценарий VANET [8].

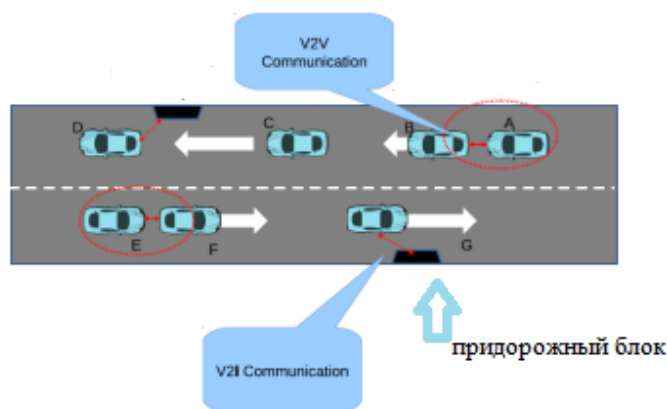


Рис. 4. Автомобильные беспроводные сети Vehicular ad hoc network (VANET)

2.5. Специальная сеть летающих объектов - Flying Ad Hoc Network (FANET) на примере сети для беспилотных летательных аппаратов

FANET можно определить как новую форму MANET, в которой узлами являются беспилотные летательные аппараты (БПЛА). Согласно этому определению, системы с одним БПЛА не могут образовывать FANET, которая может быть организована только для систем с несколькими БПЛА (см. рис.5). С другой стороны, не все системы с несколькими БПЛА образуют FANET. Связь БПЛА должна быть реализована с помощью специальной Ad Hoc сети между БПЛА. Следует указать, если связь между БПЛА полностью реализуется на связи между БПЛА с использованием инфраструктуры, то ее нельзя классифицировать как FANET.

FANET связь основана на каналах передачи данных между БПЛА и БПЛА, а не на каналах передачи данных между БПЛА и инфраструктурой. За счет взаимодействия между беспилотными аппаратами расширяется зона охвата действия связи. Даже если узел FANET не может установить канал связи с инфраструктурой, он все равно может работать, взаимодействуя через другие БПЛА. Сценарии видов такой связи показаны на рис. 5 [9].

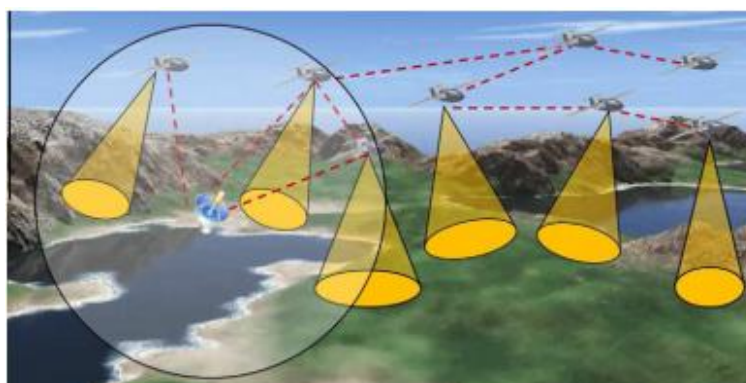


Рис. 5. FANET сценарий

Узлы этих сетей обладают способностью самим находить друг друга и формировать сеть, а в случае выхода из строя какого-либо узла могут устанавливать новые маршруты для передачи сообщений (данных).

3. Комментарии к алгоритмике и протоколам взаимодействия устройств в Ad-hoc сетях

Связь между двумя или более устройствами - сложный процесс, поэтому концептуально его необходимо разделить на несколько уровней. Он основывается на модели взаимодействия открытых систем (модель OSI), которая определяет многоуровневую архитектуру для сети передачи данных. Уровни модели OSI (MobileMan) показаны на рис 6.

Основываясь на перспективе увеличения скорости передачи данных, можно сделать вывод, что она должна быть «вертикальной» за счет одновременного улучшения нескольких уровней OSI и использования межуровневого взаимодействия. Разработка новых стандартов нижнего уровня нежелательна, поскольку Ad-Нос сеть построена на широко распространенном оборудовании стандарта IEEE 802.11. Главный вклад должен заключаться в улучшении и эффективном применении существующих протоколов. Особо следует отметить на повышение возможностей сетевого и прикладного уровня модели OSI [10].

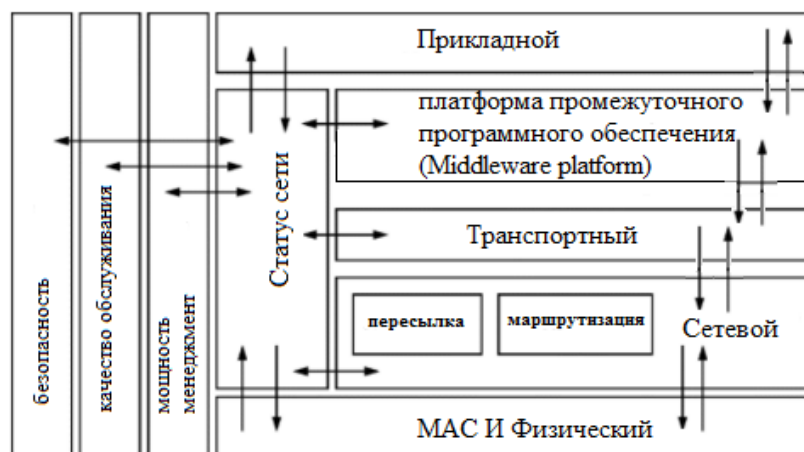


Рис. 6. Уровни модели OSI (MobileMan)

Стандарт IEEE 802.11 определяет механизмы передачи только для физического уровня OSI и подуровня MAC канала данных. Ad-Нос сети основаны на сетевом уровне, который позволяет устройствам самоорганизовываться в сети и отправлять данные во многих переходных уровнях (см. рис.6). Функциональность сетевого уровня отличает специализированную сеть от беспроводных сетей инфраструктурного типа, что делает ее новым типом сети.

Множественные скачкообразные передачи позволяют передавать данные узлам, если данные не могут быть получены напрямую из-за ограниченного расстояния передачи (радиосвязи). Для этого требуется протокол маршрутизации, чтобы гарантировать доставку пакета данных маршрутизации к месту назначения. Радиоканал используется для передачи на физическом уровне, поэтому он подвержен помехам, которые ограничивают скорость передачи данных и могут вызвать потерю пакетов. Поэтому выбор параметров формирования и обработки радиосигналов должен основываться на их помехоустойчивости.

4. Сравнительная таблица параметров современных видов самоорганизующихся сетей Ad-Нос

Для сопоставления современных видов самоорганизующихся сетей Ad-Нос в табл. 1 приводятся их основные параметры и характеристики.

Таблица 1

Сравнительная таблица параметров современных видов самоорганизующихся сетей Ad-Нос

Класс (вид) системы	Диапазон радиочастот (ГГц)	Виды модуляций радиосигналов	Дальность (М)	Мобильность	Возможности взаимодействия с другими системами и сетями	Изменение топологии / расстояние между узлами	Виды применения
MANET IEEE 802.11	2,4- 5	OFDM BPSK QPSK 16 QAM 64 QAM	До 100	низкая	Да-интернет, Wi-Fi, сотовые сети связи	медленное изменение/ короткое расстояние	связь во время боевых действий
WSN IEEE 802.15.4	2,4	DSSS QPSK ASK	До 75	очень низкая	Да-интернет, Wi-Fi, сотовые сети связи	Очень медленное изменение/ короткое расстояние	Наблю дение / сбор данных
WMN IEEE 802.11	2,4- 5	OFDM BPSK	До 100	низкая	Да-интернет, Wi-Fi, сотовые сети связи	Медленное изменение/ короткое расстояние	Военн ая связь и развед ка
VANET IEEE 802.11p DSRC	5,9	QPSK 16QAM, 64QAM BPSK	До 300	высокая	Да-интернет, Wi-Fi, сотовые сети связи	Быстрое изменение/ среднее расстояние	Помо щь водите лю/ин форми рованы е
FANET IEEE 802.11 IEEE 802.16	2,4- 11	OFDM BPSK QPSK 16 QAM 64 QAM	От 100 До 10000	очень высокая	Да-интернет, Wi-Fi, сотовые сети связи	Очень быстрое изменение/ большое расстояние	Наблю дение / сбор данных

Выводы

Проведенный сравнительный анализ, классификация и оценка вариантов реализации самоорганизующихся Ad-Нос сетей в сравнении с другими видами подвижной радиосвязи позволяет обоснованно подойти к основным принципам выбора того или иного вида подвижной радиосвязи в конкретных условиях и заданных требованиях к ее функциональным возможностям.

Библиографический список

1. Бельфер Р.А., Угрозы информационной безопасности в беспроводных саморегулирующихся сетях, 2011.
2. Lenders V., Wagner, J., Heimlicher, S., May, M. and Plattner, B. (2008); An Empirical Study of the Impact of Mobility on Link Failures in an 802.11 Ad Hoc Network, IEEE Wireless Communications, pp 16 – 21.
3. Mobile Ad Hoc Network (MANET), saching.com: www.saching.com/article/manet.
4. Молчанов Д.А., Кучерявый Е.А.. Приложения беспроводных сенсорных сетей // Электросвязь. 2009. №1, С.20-22.
5. Ian F. Akyildiz, Xudong Wang. Wireless Mesh Network, «John Wiley & Sons» Ltd, 2009, P.327.

6. Status of the Dedicated Short-Range Communications Technology and Applications Report to Congress www.its.dot.gov/index.htm, Final Report—July 2015 FHWA-JPO-15-218.
7. Sabih ur Rehman, M. Arif Khan, Tanveer A. Zia, Lihong Zheng, Vehicular Ad-Hoc Networks (VANETs) - An Overview and Challenges, *Journal of Wireless Networking and Communications* 2013, 3(3): 29-38 DOI: 10.5923/j.jwnc.20130303.02.
8. Ilker Bekmezci, Ozgur Koray Sahingoz, Samil Temel, Flying Ad-Hoc Networks (FANETs): A survey, *Survey Paper*, 2013.
9. Lauris. Cikovskis, Improving data transmission efficiency in wireless Ad-hoc network, 2016.

H.M. Alshubaki, Yu.M. Tulyakov

**ANALYTICAL COMPARATIVE EVALUATION OF MODERN
KINDS OF WIRELESS SELF-ORGANIZING AD-HOC NETWORKS**

Nizhny Novgorod State Technical University n. a. R.E. Alekseev

A comparatively generalized assessment of the variants of self-organizing wireless communication networks and Ad-Hoc data transmission, which are decentralized wireless dynamic (not having a permanent structure) networks, is carried out. The connection of client (subscriber) devices is carried out from the conditions of the possibilities (due to the limited distance) of radio communication between these devices or through communication nodes, thereby creating a self-organizing network. Based on this assessment, the classification of such networks in comparison with other types of mobile communications is justified.

Keywords: Ad-Hoc Network, SON Self-Organizing Network, MANET Mobile Ad Hoc Networks, WSN Wireless Sensor Networks, WMN Wireless Mesh Networks, VANET Automotive Wireless Networks, FANET Special Flying Object Network, OSI Network Model, DSRC Short Range Radio, V2V Vehicle - Car, V2I Car - Road infrastructure, UAV Unmanned aerial vehicle.

Н.Н. Акимов¹, А.Ю. Павлин¹, В.Р. Милов², А.Ю. Кербенева², Л.А. Васильева²

РАЗРАБОТКА ИНФОРМАЦИОННОЙ СИСТЕМЫ ПОДДЕРЖКИ ПРИНЯТИЯ РЕШЕНИЙ ДЛЯ ОБЕСПЕЧЕНИЯ КИБЕРБЕЗОПАСНОСТИ СВУ АСУ ТП

¹ Филиал РФЯЦ-ВНИИЭФ «НИИИС им. Ю.Е. Седакова»

² Нижегородский государственный технический университет им. Р.Е. Алексеева

Работа посвящена проектированию и разработке информационной системы поддержки принятия решений для обеспечения кибербезопасности СВУ АСУ ТП. Разрабатываемая информационная система предназначена для помощи разработчикам СВУ при обеспечении требуемого уровня защищенности АСУ ТП в условиях ограниченных затрат.

Ключевые слова: поддержка принятия решений, кибербезопасность, АСУ ТП, АЭС.

Проблемой системного анализа аспектов кибербезопасности АСУ ТП АЭС является наличие уязвимостей АСУ ТП, которые могут приводить к ущербу при реализации угроз кибербезопасности или вследствие киберинцидентов. Цель анализа аспектов кибербезопасности АСУ ТП заключается в обеспечении безопасного и защищённого функционирования АСУ ТП в условиях возможных кибератак.

Проблема обеспечения защищённости СВУ АСУ ТП АЭС – наличие множества неопределенностей, связанных с внешними угрозами и надежностью оценок достаточности мер защиты при недостатке (практически полном отсутствии статистических данных) кардинально затрудняет обоснование того, что требуемый уровень защищенности АСУ ТП достигнут без многократного превышения затрат на реализацию комплекса мер кибербезопасности.

Применительно к системному анализу кибербезопасности АСУ ТП АЭС поддержка принятия решения при проектировании системы кибербезопасности заключается в определении такой конфигурации используемых средств защиты, которая обеспечивает покрытие всего множества мер кибербезопасности АСУ ТП АЭС, при минимальной суммарной стоимости такой конфигурации.

Формирование модели кибербезопасности СВУ АСУ ТП АЭС заключается в формировании матрицы покрытия, включающей в себя множество мер, которые должны быть покрыты средствами, выбираемыми проектировщиком из занесенных в СППР средств.

Разработка способа принятия решений заключается в поиске и описании подхода к решению задачи выбора минимально необходимого подмножества мер обеспечения защищенности в соответствии с выбранным критерием, основанным на показателях затрат, защищенности и функциональности СВУ АСУ ТП АЭС. В достаточно общем случае подход к принятию решений по выбору мер обеспечения защищенности может быть формализован, с помощью математической постановки задачи о покрытии множества. При этом в зависимости от используемых предположений и соответствующей формализации целевых функций и ограничений возможны различные варианты постановки задачи о покрытии множества (данная задача обобщает NP-полную задачу о вершинном покрытии и потому является NP-сложной).

Архитектура информационной системы поддержки принятия решений для обеспечения кибербезопасности СВУ АСУ ТП АЭС (ИС ППР КБ) представляет собой базовую организацию системы, состав компонентов и их отношения между собой и с окружением, а также принципы, определяющие проектирование и развитие системы.

Архитектура ИС ППР КБ представлена на рис. 1.

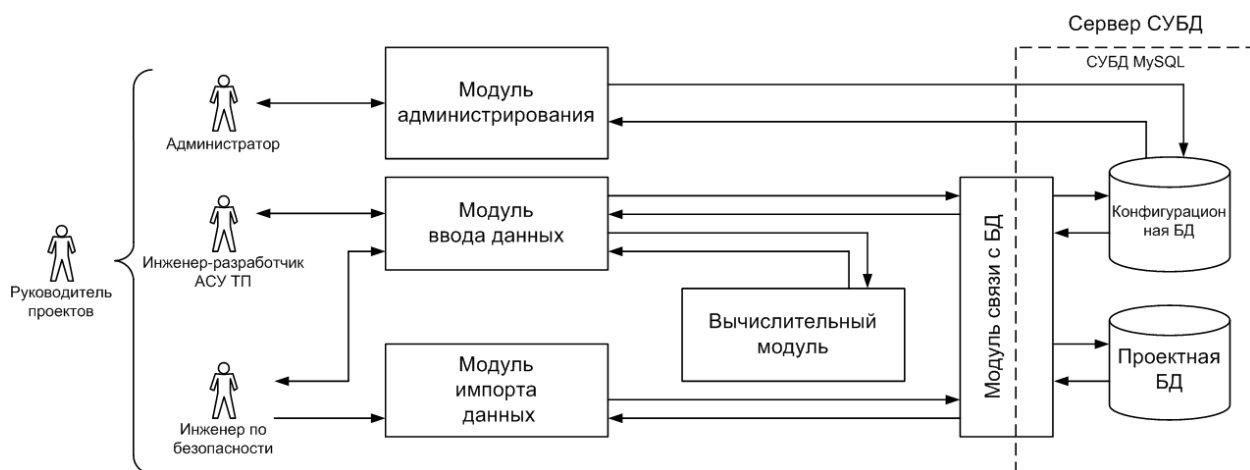


Рис. 1. Архитектура ИС ППР КБ

Алгоритм работы ИС ППР КБ представлен в виде блок-схемы на рисунке 2.



Рис. 2. Блок-схема алгоритма работы ИС ППР КБ

В качестве входных данных используется задание пользователя на определение конфигурации средств защиты для конкретного проекта по разработке АСУ ТП, которое формирует пользователь в роли «Разработчик АСУ ТП», а также список средств защиты, который формирует пользователь в роли «Инженер по безопасности».

В качестве управляющей информации выступает национальная нормативная база, которая включает в себя меры сопоставленные друг с другом безопасности и классы защищенности по приказам ФСТЭК №31 от 14 марта 2014 г. и № 239 от 25.12.2017 г.

На выходе пользователь в роли «Разработчик АСУ ТП» может получить решения по конфигурации средств защиты для покрытия всех мер по выбранному приказу ФСТЭК для конкретного проекта, сохранить лучшую конфигурацию, а также просмотреть матрицу «Проекты-Средства защиты» для формирования спецификации средств защиты для конкретного проекта.

На начальном этапе работы алгоритма поддержки принятия решений указываются входные данные. В качестве таких данных могут выступать, например, меры обеспечения защищенности (согласно приказам ФСТЭК № 31 и № 239), параметрами являются стоимость средств и результативность их применения.

Согласно требованиям, определяются m необходимых для покрытия мер и n возможных к использованию средств защиты. Таким образом, формируется матрица «меры-средства» D размерностью $m \times n$, элементами которой являются 0 или 1. Если j -ое средство покрывает i -ую меру, то элемент матрицы d_{ij} принимает значение равное 1. В противном случае – 0.

$$D = \begin{bmatrix} d_{11} & d_{12} & \dots & d_{1n} \\ d_{21} & d_{22} & \dots & d_{2n} \\ \dots & \dots & \dots & \dots \\ d_{m1} & d_{m2} & \dots & d_{mn} \end{bmatrix}, \text{ где } i = \overline{1, m}; j = \overline{1, n}$$

Определим k -ю конфигурацию средств защиты, как вектор $\mathbf{v}_k = [v_{1k}, \dots, v_{Nk}]^T$, где N – количество всех доступных для применения средств защиты. Компоненты вектора конфигурации средств $v_{nk} = 1$, если n -е средство используется в k -й конфигурации и $v_{nk} = 0$ в противном случае. Каждое средство защиты характеризуется стоимостью c_n , $n = \overline{1, N}$. Тогда стоимость совокупности средств защиты, соответствующих k -й конфигурации определяется выражением

$$C_k = \sum_{n=1}^N v_{nk} c_n = \mathbf{c}^T \mathbf{v}_k$$

Здесь $\mathbf{c} = [c_1, \dots, c_N]^T$ – вектор стоимостей отдельных средств защиты. Каждое средство защиты покрывает одну или несколько мер защиты, что представляется вектором $d_n = [d_{1n}, \dots, d_{Mn}]^T$. Все доступные для применения средства защиты могут быть охарактеризованы $(M \times N)$ -матрицей \mathbf{D} «меры-средства». Элементы этой матрицы $d_{mn} = 1$, если n -е средство покрывает m -ю меру и $d_{mn} = 0$ в противном случае.

Множество допустимых конфигураций средств защиты составляют те меры, которые обеспечивают защищенность объекта защиты. Условие защищенности является выполненным при наличии в k -й конфигурации средства защиты для каждой меры. Для проверки условия защищенности введем вектор

$$\mathbf{a}_k = \mathbf{D} \mathbf{v}_k,$$

компоненты которого равны количеству средств, покрывающих соответствующие меры. Если для k -й конфигурации средств $a_{mk} \geq 1$, $m = \overline{1, M}$, то защищенность АСУ ТП будет обеспечена. Представим условие защищенности в виде $s_k \geq 1$, где

$$s_k = \min_{m=1, M} a_{mk}$$

Условие защищенности будет выполняться в случае, когда все компоненты вектора $a_{mk} = 1$, $m = \overline{1, M}$. При этом каждая мера покрывается одним средством, что является

достаточным для обеспечения защищенности. Однако на практике ряд мер могут быть покрыты несколькими средствами, что приводит к избыточности полученного решения.

Пример пользовательского интерфейса для просмотра проекта представлен на рис. 3. Из проекта возможно перейти в матрицу «Меры – средства защиты», представленной на рис. 4.

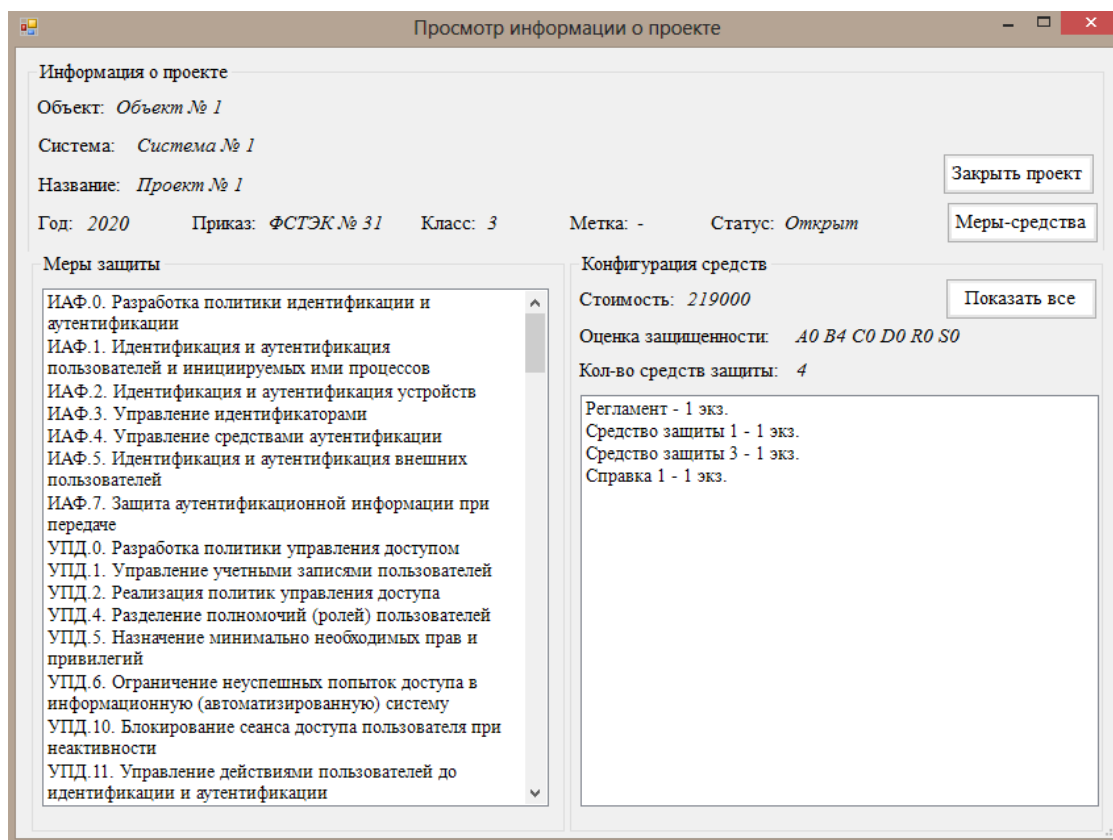


Рис. 3. Просмотр информации о проекте в базе данных

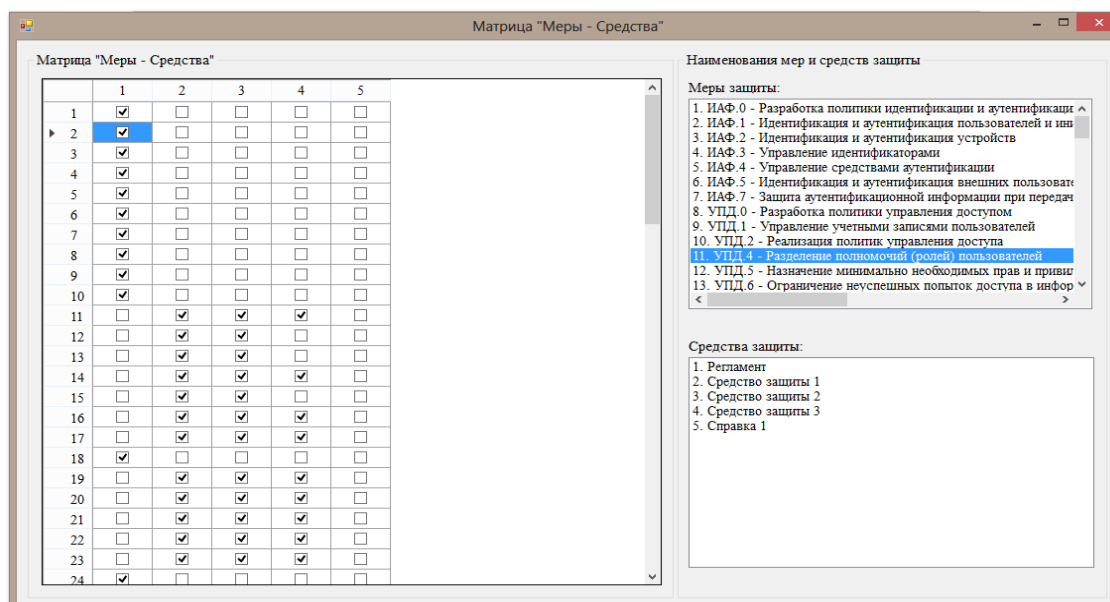


Рис. 4. Просмотр матрицы «Меры – средства защиты»

Основной процесс работы с программой заключается в том, чтобы в окне для покрытия мер средствами защиты (рис. 5) пользователь в роли «Инженер по безопасности» для каждой меры выбрал необходимые средства защиты, а также указал их количество и цену.

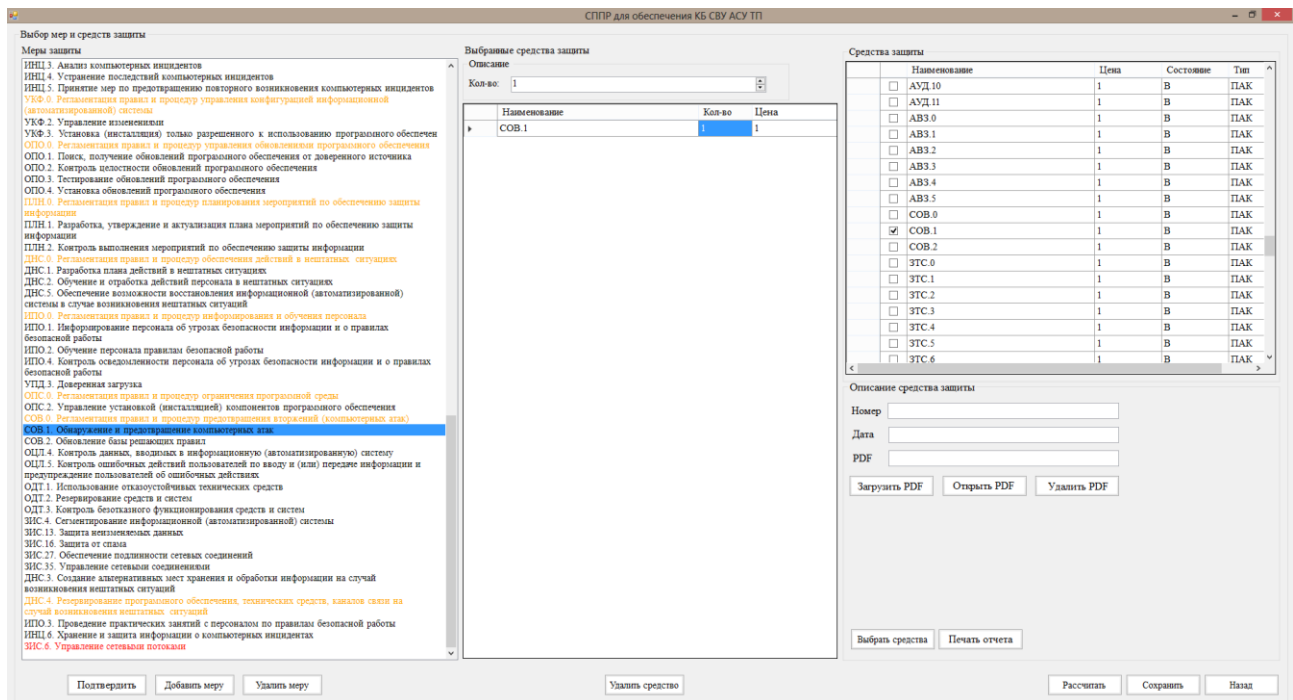


Рис. 5. Окно для покрытия мер средствами защиты

Пользователь в роли «Разработчик АСУ ТП» может запустить расчет конфигураций только после того, как все меры покрыты средствами защиты.

Таким образом, поддержка принятия решений при создании комплексной системы кибербезопасности СВУ АСУ ТП АЭС позволит покрыть все требуемые меры безопасности, обозначенные в нормативной документации, необходимыми средствами защиты при минимальной стоимости или максимальной защищенности.

Автоматизация процесса покрытия мер безопасности будет способствовать сокращению сроков реализации проектов, минимизации ошибок при выборе средств защиты, а также повышению информированности ЛПР (в части формирования отчетов).

Библиографический список

1. Дмитриев С.М., Акимов Н.Н., Кольцов В.А., Аспекты обеспечения кибербезопасности АСУ ТП АЭС // Информационно-измерительные и управляющие системы. 2017. № 8. С. 7–13.
2. Акимов Н.Н., Милов В.Р., Егоров Ю.С. Элементы концептуальной модели обеспечения кибербезопасности критически важных промышленных объектов // Материалы XXIV Международной научно-технической конференции «Информационные системы и технологии» (ИСТ – 2018). – Н. Новгород: НГТУ, 2018. С. 463–467.
3. Методы проектирования информационно-управляющих и телекоммуникационных систем / Под ред. В.Р. Милова, В.Г. Баранова. – М.: Радиотехника, 2016. – 216 с.
4. Леончик П.В. Алгоритм поиска покрытия разреженных булевых матриц // Информатика. 2007. № 2. С. 53 – 61.

N.N. Akimov¹, V.A. Koltsov¹, A.Y. Pavlin¹, V.R. Milov², A.Y. Kerbeneva², L.A. Vasilyeva²

**DEVELOPMENT OF DECISION SUPPORT INFORMATION SYSTEM FOR
ENSURING CYBERSECURITY OF NUCLEAR POWER PLANT INDUSTRIAL
CONTROL SYSTEM**

¹ Branch of «Russian Federal Nuclear Center – All-Russian Scientific Research Institute of
Experimental Physics» (RFNC-VNIIEF) «NIIS named after Yu. Ye. Sedakov»

² Nizhny Novgorod State Technical University n.a. R.E. Alekseev

The work is devoted to the design and development of an information system for decision support to ensure cyber security nuclear power plant (NPP) industrial control system (ICS). The developed information system is designed to help the developers of the ICS in ensuring the required level of security under conditions of limited costs.

Keywords: decision support, cyber security, Industrial control system (ICS), Nuclear power station.

ПРЕДИКТОР СОСТОЯНИЙ СТЕКА ПРОТОКОЛОВ FANET

Нижегородский государственный технический университет им. Р.Е. Алексеева

Обсуждается интегрированное приложение «Предиктор состояний стека протоколов FANET» когнитивной радиостанции (CR), являющейся узлом авиационной AdHoc сети беспилотных летательных аппаратов (БПЛА) (FANET - Flying Ad Hoc Network). Это приложение оценивает и прогнозирует состояния стека протоколов летающей самонастраивающейся сети на базе модели «пространства переменных состояний» (VSS – Variable State Space) и фильтра Калмана [1].

Ключевые слова: предиктор состояний стека протоколов FANET, модель пространства переменных состояний (VSS – VariableState Space); настройка диаграмм направленности антенн, адаптация радиолинии, компенсация Доплера и задержки распространения сигнала.

Важнейшим назначением сетей FANET является надёжная и своевременная передача данных от источника к получателю. Однако, эта задача остаётся вызывающей проблемой из-за высоко динамичной природы летающих сетей, которым свойственны скорости движения до 3600 км/с, дальности прямой радиовидимости до 400 – 800 км, зеркальное отражение луча от гладкой земной поверхности из-за высокой поднятости антенн. При этом имеют место большие доплеровские сдвиги частоты и фазы сигналов, большие задержки и потери распространения сигналов в свободном пространстве, что существенно снижает эффективность протоколов множественного доступа к каналу с временным разделением (TDMA), требует реализации мер по коррекции Доплера, задержек и потерь распространения сигнала в свободном пространстве, межсимвольной интерференции для обеспечения высоких пропускных способностей и надёжности сети связи. Компенсация Доплера и времени задержки распространения должна осуществляться с помощью прогнозирования этих параметров на каждом узле и предсказания передаваемых сигналов по частоте и времени начала передачи пакета на величину, обратную по знаку прогнозируемым значениям, чтобы сигнал приходил на вход приёмника в идеальном виде – без Доплера и без задержки, т.е. в начале слота. Именно такая технология компенсации Доплера и задержки распространения сигнала реализуется в новой системе связи «Воздух-Земля» и «Воздух-Воздух» LDACS-1 гражданской авиации [2-3]. Потери распространения сигнала в свободном пространстве (пропорциональные квадрату расстояния и квадрату номинала частоты) должны компенсироваться высокими коэффициентами усиления остронаправленных антенн. Для точной настройки диаграмм направленности антенн необходимо прогнозировать 3D координаты, 3D вектор скорости и 3D позиции корпуса ЛА передатчика и приёмника на момент времени прихода сигнала на вход приёмника.

Кроме высокой динамики и большой дальности трассы, другой ключевой причиной проблемы низкой надёжности авиационной связи является отсутствие обратной связи нижних уровней эталонной модели взаимодействия открытых систем (ЭМОС) (физического, канального, сетевого) с верхними уровнями (транспортным, сеансовым, представления, приложения). При этом верхние уровни пытаются гарантировать требуемую вероятность доставки пакетов с помощью процедур автоматического запроса повторения (ARQ – Automatic Repeat Request), что в каналах с низким качеством линии приводит к обратному эффекту - к большому числу ретрансляций, большим задержкам, потерям пакетов, разрыву линии. TCP хорошо работает в сетях MANET, но недопустим для FANET, где требуется протокол UDP (без ARQ), не простой UDP, а надёжный [4] с мощным каскадным кодом, исправляющим большие пакеты ошибок.

Чтобы решить проблему низкой эффективности транспортного и прикладного уровней при низком качестве нижних уровней стека протоколов, предлагается прогнозировать все существенные параметры качества стека протоколов авиационной сети, объединить их в единый вектор параметров, и прогнозировать состояния этого вектора с помощью специального агрегированного приложения (предиктора состояний стека протоколов). Предиктор должен иметь программируемый интерфейс с операционной системой (ОС), соответствующий интерфейсу API (Application Programming Interface), заданному в ARINC 653 [5] для интегрированной модульной авионики (ИМА). Операционная система должна быть многозадачной, реального времени (RTOS), не хуже, чем VxWorks 653, а платформа для разработки не хуже, чем Workbench for VxWorks 653 и другие инструменты проектирования ПО из состава пакета VxWorks 653 Platform [5], который сертифицирован согласно руководству по сертификации программного обеспечения авионики (документу DO-178B, C) [6]. Интерфейс API работает, как трёхуровневый стек (приложение – интерфейс - ОС) автономный для каждого раздела приложения. Он исключает взаимные влияния отказов разных разделов приложений, предоставляя каждому разделу приложения своё пространство памяти, свой раздел ОС, свои временные слоты для взаимодействия между разделом приложения и ОС. API также называют протоколом APEx (Application – Execution). Он позволяет одно и то же приложение использовать для разных ОС – переносить приложение на разные исполнительные платформы.

Предиктор состояний стека протоколов через программируемый интерфейс API должен постоянно предоставлять каждому уровню (протоколу) стека FANET вычисленный им вектор параметров (состояний), чтобы каждый уровень мог принимать корректные решения по своим оптимальным режимам работы с учётом состояний всех других уровней ЭМВОС. Для прогнозирования состояний параметров стека протоколов целесообразно использовать модель «пространства переменных состояний» (VSS – Variable State Space) [1, 7], которая прогнозирует состояния параметров качества всех уровней стека протоколов (канала, линии, сети, транспорта, сеанса, представления, приложения) с помощью фильтра Калмана, гарантирующего минимум квадрата ошибки слежения, и интерфейс API между всеми приложениями радиостанции и операционной системой, который предоставляет прогнозы этих параметров для всех перечисленных уровней когнитивной радиостанции (CR).

Отслеживаемыми параметрами являются: 1) на физическом уровне: доплеровские сдвиги частоты и фазы, задержка распространения сигнала в свободном пространстве (для компенсации этих параметров путём предискажения передаваемого сигнала), индикатор уровня принимаемого сигнала (RSSI - Received Signal Strength Indicator) и отношение «сигнал/шум» (SNR - Signal-to-Noise Ratio) (для адаптации по виду сигнально-кодовой конструкции (СКК)); координаты 4D траектории (широта, долгота, высота, 3D вектор скорости) и 3D позиции корпуса своего ЛА и ЛА соседа - получателя пакета (углы азимута, тангажа, крена) (для оценки дальности трассы с целью расчёта потерь распространения сигнала для точной настройки диаграмм направленности антенн); 2) на уровне линии: показатель качества линии (LQI - Link Quality Indicator) для многопараметрической адаптации радиолинии; 3) на уровне сети: вероятность правильного приёма пакета (PRR – Packet Reception Rate), задержка передачи пакета (для выбора маршрута); 4) на транспортном уровне: показатель L перегрузки сети, приращение длины очереди dL, вероятность ошибочного приёма пакета (для адаптации размера окна квитирования, протокола доступа к каналу, избыточности каскадного кода).

Идея объединения показателей качества разных уровней модели ЭМВОС в один вектор основана на предпосылке о том, что для достижения лучших характеристик сети FANET в целом уровень приложения должен взять на себя ответственность управления всем стеком протоколов FANET с учётом всех параметров: физического, канального, сетевого, транспортного, представительного, прикладного уровней, чтобы

соответствующим образом корректировать своё поведение, например, менять избыточность каскадного кодирования пакетов на уровне приложения (или транспорта) и исправлять ошибки, обнаруженные на транспортном уровне, вместо того, чтобы передавать повторно пакет с помощью ARQ по плохому физическому каналу, что приводит к потере пакетов и самой радиолинии из-за переполнения очереди к каналу [4].

В настоящее время большинство подходов для оценки качества линии и её адаптации основаны на таких вышеупомянутых метриках как PRR, LQI, SNR и RSSI.

Однако, существующие подходы оценки этих метрик ограничены в своей способности предсказывать будущее - стохастическое поведение динамических параметров качества линии и сети. Они [7] используют модель авторегрессии (AR) первого порядка для вычисления отношения сигнал/шум (SNR). Однако эта модель слишком проста, так как является скалярной моделью и предполагает, что возмущения являются гауссовыми с известными отклонениями, что в большинстве практических случаев неверно. Для моделирования мгновенных изменений (эволюции) параметров узкополосного Марковского случайного процесса оптимален метод пространства переменных состояний, предложенный и развитый в работах Кловского Д.Д., Стратоновича Р.Л., Калмана Р., Тихонова В.И., Ярлыкова М.С., Сосулина Ю.Г. и др. учёных [1]. Применение этого подхода существенно расширило возможности оптимального синтеза адаптивных приёмников [1]. Согласно этому методу Марковские случайные процессы изменений параметров канала можно моделировать с помощью стохастических дифференциальных уравнений (СДУ) первого порядка:

$$\frac{dx(t)}{dt} = \dot{x}(t) = -\alpha x(t) + n_x(t) \quad (1)$$

Это уравнение называется уравнением состояния переменной $x(t)$ или уравнением формирующего фильтра состояния переменной $x(t)$. Здесь $n_x(t)$ – формирующий (порождающий) белый гауссовский нормированный шум модели состояния переменной $x(t)$ с нулевым математическим ожиданием $M[n_x(t)] = 0$ и единичной дисперсией $D_{n_x} = 1$, α – коэффициент, характеризующий скорость изменения случайного процесса $x(t)$, равный ширине спектра процесса.

Поскольку обработка сигналов в приёмнике осуществляется в цифровом виде с использованием отсчётов в дискретные моменты времени, от дифференциальных уравнений состояния (1) переходят к разностным уравнениям состояния. Для этого уравнение (1) умножают на дискрет времени dt , равный Δt – шагу дискретизации по времени, который выбирается согласно теореме Котельникова, т.е. частота дискретизации

F_d берётся не менее удвоенной ширины полосы моделируемого процесса $F_d = \frac{1}{\Delta t} \geq$

2α , отсюда $\Delta t \leq \frac{1}{2\alpha}$. Также в (1) заменяют приращение $dx(t)$ на разность смежных дискретных отсчётов $x_k - x_{k-1}$ и получают разностное уравнение

$$\begin{aligned} x_k - x_{k-1} &= -\alpha \Delta t x_{k-1} + \Delta t n_{k-1}, \text{ откуда} \\ x_k &= (1 - \alpha \Delta t) x_{k-1} + \Delta t n_{k-1}. \end{aligned} \quad (2)$$

Выражение $(1 - \alpha \Delta t)$ при малых значениях $\alpha \Delta t$ можно заменить на $e^{-\alpha \Delta t} = R(\tau)$ – коэффициент корреляции Марковского процесса между смежными отсчётами. Отсюда разностное уравнение состояния переменной x_k записывается:

$$x_k = e^{-\alpha \Delta t} x_{k-1} + n_k = R(\Delta t) x_{k-1} + n_k. \quad (3)$$

Здесь $R(\Delta t)$ – коэффициент корреляции или вероятность перехода переменной из одного состояния x_{k-1} в следующее x_k за время $\tau = \Delta t$. В случае векторного (многокомпонентного) процесса обозначим его \bar{x} . При этом коэффициент корреляции

превращается в системную матрицу $\bar{\Phi}$ (матрицу вероятностей переходов вектора переменных из состояния \bar{x}_{k-1} в состояние \bar{x}_k за один шаг Δt). Разностное уравнение состояния вектора переменных \bar{x} имеет вид

$$\bar{x}_k = \bar{\Phi} \bar{x}_{k-1} + \bar{\Gamma} \bar{n}_k, \quad (4)$$

где k – номер временного отсчёта в момент времени $t_k = k\Delta t$, $\bar{\Gamma}$ – матрица «вход - состояние» или шумовая матрица, преобразующая входной (порождающий) шум модели в вектор состояния. Функциональная схема дискретной модели состояния вектора параметров \bar{x} представлена на рис. 1 [1].

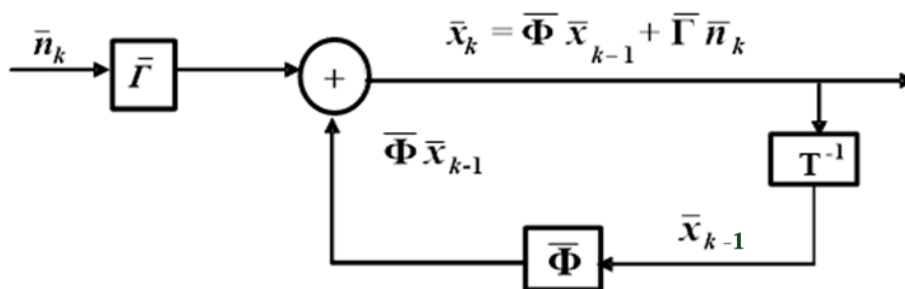


Рис. 1. Дискретная модель состояния векторной переменной \bar{x}

Для линейной модели состояния и линейной модели наблюдения оптимальным устройством оценивания и прогнозирования (фильтрации) вектора переменных параметров стека протоколов \bar{x} , обеспечивающим минимум среднеквадратической ошибки фильтрации, является расширенный (векторный) фильтр Калмана, структурная схема которого представлена на рис. 2 [1]. Этот фильтр построен на базе дискретной модели состояния векторной переменной, приведённой на рис. 1, и модели наблюдения, которая представляет смесь аддитивного белого гауссовского (АБГШ) наблюдения и известного обучающего сигнала, переданного через канал с неизвестными изменяющимися во времени параметрами. Правая часть фильтра Калмана является моделью состояния вектора переменных параметров канала \bar{x} , представленной справа на рис.1.

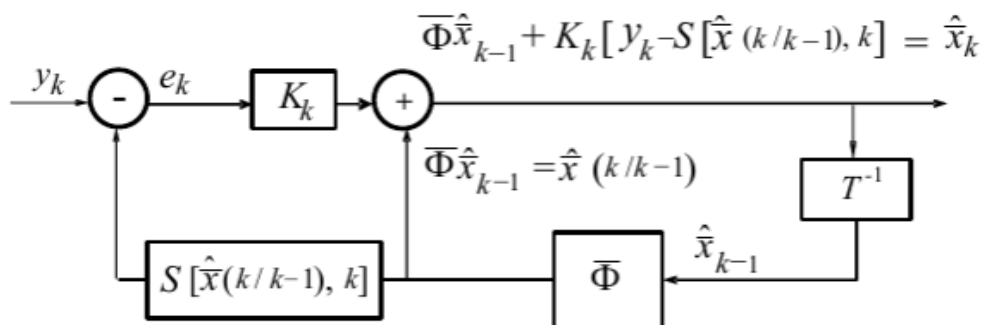


Рис. 2. Дискретный фильтр Калмана

В правой части фильтра Калмана вычисляются текущая оценка \hat{x}_k вектора параметров \bar{x} , а также прогнозируемая (априорная) оценка $\bar{\Phi} \hat{x}_{k-1}$ вектора параметров канала, которая зависит от системной матрицы состояния и предыдущей оценки и является первым слагаемым текущей оценки \hat{x}_k . Кроме того, прогнозируемая оценка параметра поступает

в левую часть фильтра, где она подставляется в формулу прогнозируемой оценки отсчёта принимаемого сигнала $S[\hat{\bar{x}}_k(k/k - 1)]$.

В левой части фильтра вычисляется ошибка $e(k)$ прогнозируемого отсчёта сигнала путём сравнения его с реализацией смеси сигнала и шума $y(k)$ на входе приёмника и вычисляется оптимальный весовой коэффициент K_k фильтра Калмана. Ошибка прогноза, умноженная на коэффициент Калмана, складывается с прогнозируемой оценкой параметров канала, обеспечивая текущую оценку вектора параметров $\hat{\bar{x}}_k$. Фильтр Калмана гарантирует минимальный средний квадрат ошибки прогнозирования параметров качества стека протоколов. На рис. 3 приведён пример прогнозирования индикатора качества линии LQI (Link Quality Indicator) фильтром Калмана. На рис. 4 приведён пример прогнозирования LQI на базе модели авторегрессии AR [7].

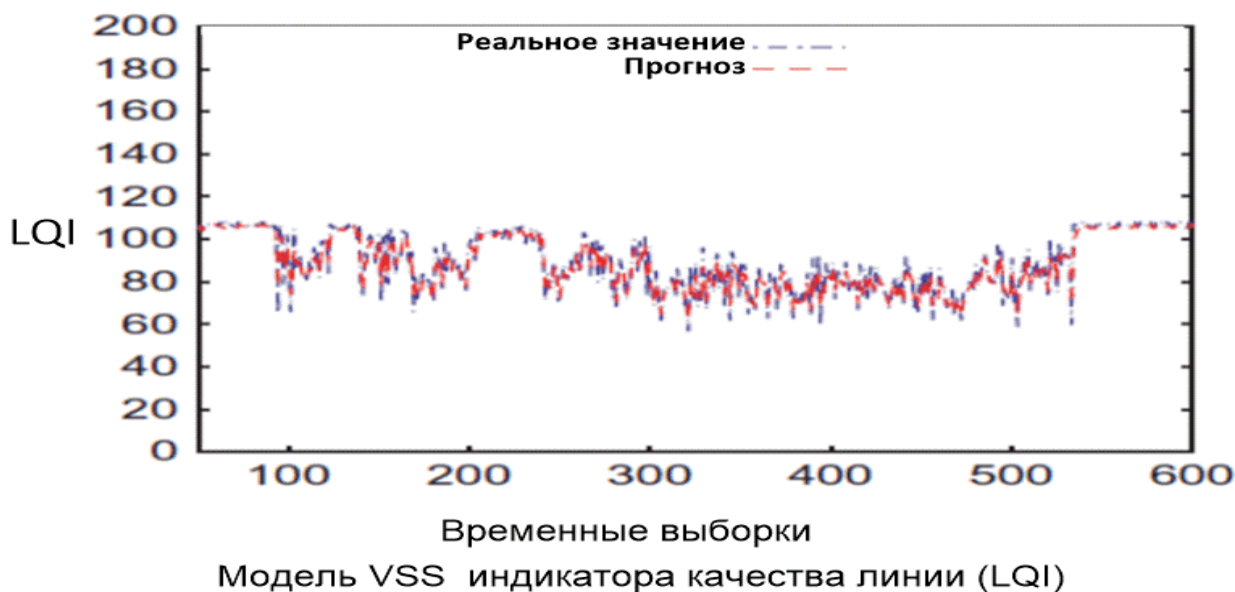


Рис. 3. Пример прогнозирования индикатора качества линии LQI на базе модели VSS (фильтром Калмана)

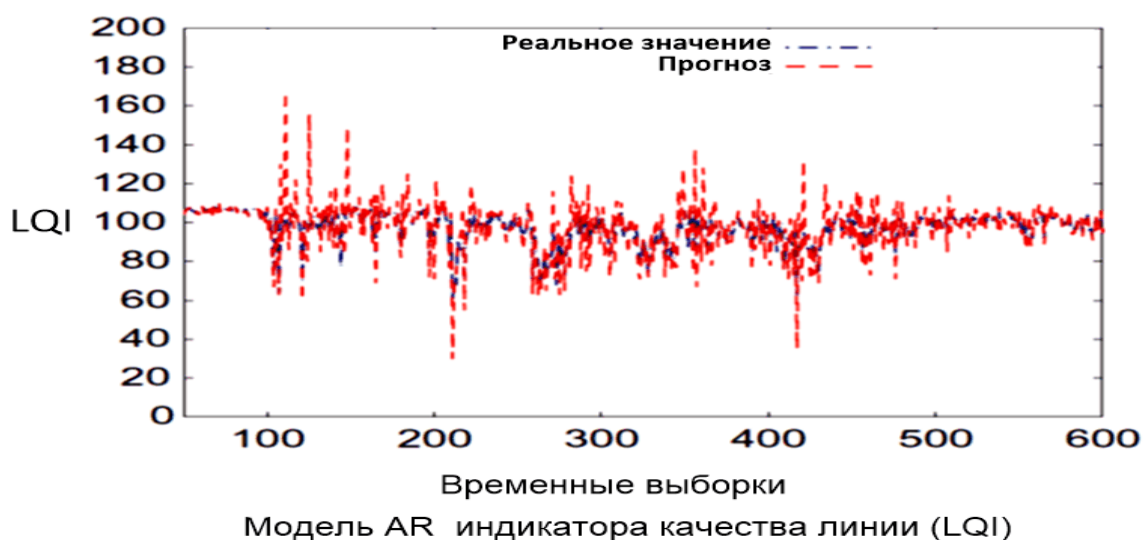


Рис. 4. Пример прогнозирования индикатора качества линии (LQI) на базе модели авторегрессии AR

Из рис. 3 и рис. 4 видно, что фильтр Калмана достаточно плавно отслеживает и сглаживает изменения индикатора качества линии (LQI) в то время, как фильтр авторегрессии (AR) вычисляет много всплесков, не имеющих основания.

В данной статье представлены общая концепция, математическая модель, блок-схема и результаты моделирования в Mathcad «Предиктора состояний стека протоколов FANET» в «пространстве переменных состояний» (VSS) на базе фильтра Калмана. Дальнейшее исследование и моделирование «Предиктора состояний стека протоколов FANET» будет продолжено с использованием языка Python и фокусированием внимания на алгоритмах многопараметрической адаптации радиолинии, радиосети и верхних уровней стека протоколов FANET.

Библиографический список

1. Горячева Т.И. Теория информации в системах связи: учеб. пособие / Т.И. Горячева, В.Р. Милов; Нижегород. гос. техн. ун-т им. Р.Е. Алексеева. – Нижний Новгород, 2014. – 247 с.
2. L-DACS1 System Definition Proposal: Deliverable D3: Design Specifications for L-DACS1 Prototype // EUROCONTROL, Tech. Rep. Version 1.0. 2009.
3. <http://www.eurocontrol.int/sesar-research/publications>
4. Комяков А.В. Характеристики радиолиний в системах авиационной электросвязи: учеб. пособие / А.В. Комяков, В.Р. Милов, Т.И. Горячева; Нижегород. гос. техн. ун-т им. Р.Е. Алексеева. – Нижний Новгород, 2018 – 222 с.
5. Q. Luo and J. Wang, “FRUDP: A reliable data transport protocol for aeronautical ad hoc networks”. DOI 10.1109/JSAC.2018.2804099, IEEE Journal on Selected Areas in Communications. 2018. Vol.36 (2). P. 257 - 267.
6. Демьянов А.В. Современные технологии и средства разработки программного обеспечения критичных по безопасности систем, сертифицируемых по требованиям DO-178B. AVD Systems. www.avdsys.ru.
7. КТ178В. Рассмотрение вопросов сертификации программного обеспечения авионики. 2011.
8. Ma X., Liao J., Djouadi S. M., Cao Q. LIPS: Link Prediction as a Service for data aggregation applications// Elsevier. Ad Hoc Networks.2014. Vol. 19. P. 43–58.

A.YU. Manakov, A.E. Nikishin, T.I. Goryacheva

PREDICTOR OF THE FANET PROTOCOL STACK STATES

Nizhny Novgorod State Technical University n.a R.E. Alekseev

An integrated application "Predictor of the FANET protocol stack states" of an aviation cognitive radio station (ACR), which is a node of the aviation Ad Hoc network of unmanned aerial vehicles (UAVs) (FANET - Flying Ad Hoc Network), is discussed. This application evaluates and predicts the state of the protocol stack of a self-adapting network based on the model of VSS (variable states space) and the Kalman filter.

Keywords: FANET protocol stack states predictor, Variable State Space model (VSS-model); tuning of the antenna pattern, adaptation of the radio link, compensation of Doppler and signal propagation delay.

ИССЛЕДОВАНИЕ СИСТЕМ СВЯЗИ НА БАЗЕ ДОМЕНОВ ПРЕОБРАЗОВАНИЙ

Нижегородский государственный технический университет им. Р.Е. Алексеева

Рассмотрены базовые особенности структур передатчиков и приёмников когнитивных радиосистем, реализующих широкополосные схемы модуляции на базе прямого и обратного быстрых преобразований Фурье (*FFT* – Fast Fourier Transform). К ним относятся сигналы мультиплексирования с помощью ортогонального частотного разделения (*OFDM* – Orthogonal Frequency Division Multiplexing), сигналы множественного доступа с кодовым разделением и многими несущими (*MC – CDMA* – Multi Carrier – Code Division multiple access), также сигналы фундаментальной модуляции (*FMW* – Fundamental Modulation Waveform), используемые в системах связи на базе доменов преобразований (*TDCS* – transform-domain communication system). Предложено реализовать алгоритмы частотной адаптации для повышения спектральной эффективности систем связи с *OFDM* и *MC – CDMA*.

Ключевые слова: мультиплексирование с ортогональным частотным разделением, множественный доступ с кодовым разделением и многими несущими, системы связи на базе домена преобразований, быстрое преобразование Фурье, частотная адаптация.

Для эффективной эксплуатации спектра в динамически изменяющейся среде целесообразно систему связи сделать приспособляющейся к изменениям среды и в то же время, гарантировать отсутствие интерференции для других пользователей канала. Такие системы называются системами когнитивного радио (*CRS* - cognitive radio system) [1]. Главные задачи «когнитивного радио» состоят в: анализе спектра радио-окружения, идентификации канала, управлении частотой, мощностью, видом передаваемой сигнально-кодовой конструкции, динамическом менеджменте спектром.

Для решения этих задач необходимо прогнозировать параметры мобильного канала и адаптировать параметры радиолинии, чтобы гарантировать заданное качество связи. Учитывая огромный потенциал когнитивных радиостанций, заключающийся в спектральной гибкости систем связи будущего, рассмотрим общие характеристики и основные различия между системами связи на базе доменов преобразований (*TDCS*) и технологий *OFDM* и *MC – CDMA* [2]. Основная идея систем *TDCS* состоит в синтезе интеллектуальной (с программным управлением) структуры сигнала в частотной области для предотвращения передачи на частотах, занятых преднамеренными или естественными помехами. Этот синтез осуществляется перед передачей пакета при формировании сигнала в передатчике, вместо традиционной компенсации интерференции, реализуемой в приёмнике.

При внедрении системы связи на базе доменов преобразований *TDCS* спектральная интерференция и наличие полезного сигнала оцениваются с помощью быстрого преобразования Фурье - основного метода спектрального оценивания. Как только полосы, содержащие интерференцию или сигналы посторонних систем связи, идентифицированы с помощью спектрального оценивания и порогового детектирования, эти занятые помехами полосы эффективно режектируются перед созданием структуры передаваемого сигнала в частотной области [3, 4, 6]. В занятых помехами (режектированных) частотных полосах сигнал для передачи не генерируется. Спектральное представление передаваемого сигнала генерируется в чистых полосах, после чего, используя обратное быстрое преобразование Фурье (*IFFT* – Inverse Fast Fourier Transform) во временной области формируется сигнал фундаментальной модуляции (*FMW* – fundamental modulation waveform), который не создаёт излучений в занятых полосах. Сигнал *FMW* не содержит энергии на спектральных позициях помех.

В результате принимаемые символы не подвергаются влиянию помех от других станций и сами не создают помех. В этом состоит особенность систем связи на базе домена преобразований *TDCS*. Причём сигналы *TDCS* являются сигналами *CDMA* с расширением спектра прямой кодовой последовательностью (*DSSS* – Direct Sequence Spread Spectrum) и скремблированием псевдослучайной последовательностью фаз для крипто- и имитозащиты [3]. На рис. 1 представлена функциональная схема генерации и передачи сигналов *TDCS*, реализующая процессы: оценки спектра, пороговой обработки спектра, режекции частотных областей сигнала, где присутствуют помехи, генерации фазскремблирующей последовательности, масштабирования, обратного преобразования Фурье (*IFFT*), записи в буфер сигнала фундаментальной модуляции $b(t)$ (*FMW* – Fundamenta Modulation Waveform).

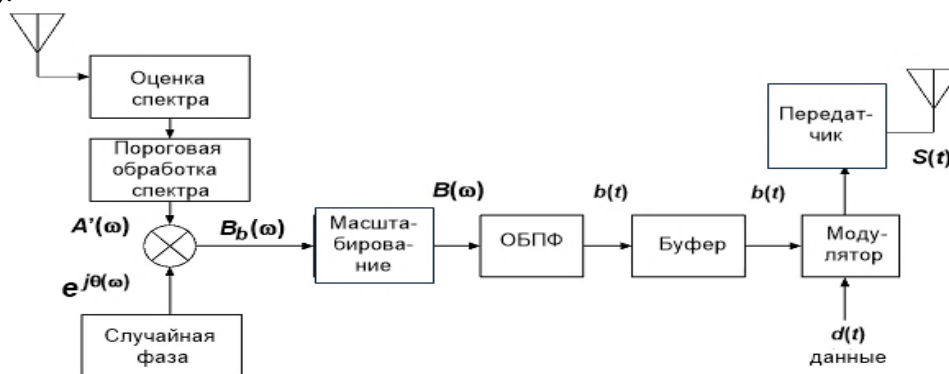


Рис. 1. Функциональная блок-схема передатчика *TDCS*

Из рис. 1 видно, что свободные от интерференции участки установлены, форма сигнала $b(t)$ фундаментальной модуляции *FMW* генерирована, сохранена, модулирована данными $d(t)$ и передана. Спектральная идентификация в передатчике на рис.1 определяет свободные от помех области спектра. Амплитуда спектра вычисляется из спектральных оценок. Для предотвращения интерференции частотных компонентов используется жёсткое пороговое ограничение. Пороговая величина обычно базируется на средней мощности, содержащейся в спектре. Применение порогов для оценки спектра генерирует «чистые» или свободные от помех области спектра $A(\omega)$. Частотные компоненты с амплитудами, превышающими порог, устанавливаются в ноль («зануляются»), а оставшимся спектральным компонентам присваивают значения 1. В блоке «Случайная фаза» (на рис.1) генерируется комплексный псевдослучайный фазовый вектор (*PR* – Pseudo Random), как скремблирующая последовательность для умножения на элементы чистого спектра $A(\omega)$ и получения спектра сигнала $B_p(\omega)$. Использование фазового вектора *PR* гарантирует, что временная реализация сигнала *FMW* имеет корреляционные свойства подобные выборкам шума. Это делается для крипто- и имитозащиты сигнала. Блок масштабирования на рис. 1 масштабирует комплексный спектр для обеспечения желаемой энергии сигнала $B(\omega)$. Это масштабирование эффективно наделяет все символы связи для передачи равными энергиями. Блок обратного быстрого преобразования Фурье (*IFFT*) (*IFFT*) в схеме передатчика *TDCS* на рис.1 генерирует фундаментальную форму волны *FMW* $b(t)$ во временной области и запоминает её в буфере. Множество символов, которые передаются с использованием одной операции *IFFT* и генерации новой формы волны *FMW* зависит от рабочих требований и изменений среды. Этот процесс отличается от традиционной техники *OFDM*, поэтому стадия *IFFT* требуется для передачи каждого символа $b(t)$ и данных $d(t)$. На рис. 2 приведены спектральные оценки, реализуемые в *TDCS*. Верхний график на рис. 2 показывает спектральную оценку репрезентативной среды, содержащей две узкополосные однотоновые помехи, которые создают соответствующий режектируемый вектор. Свободный от помех спектр приведён на нижнем графике рис. 2.



Рис. 2. Спектральные оценки и режекция полос с помехами реализуемые TDCS

Для реализации множественного доступа пользователь *TDCS* имеет в качестве адреса линейный сдвиговой регистр (*LFSR – linear feedback shift register*) с обратными связями, конфигурированный для получения на выходе двоичной последовательности максимальной длины или -последовательности. *LFSR* может быть конфигурирован для выдачи нескольких разных двоичных последовательностей, зависящих от порождающего многочлена (расположения отводов обратной связи). На рис. 3 представлены две последовательности фундаментальной модуляции *FMW*s (верхний график), генерированные с разными порождающими многочленами, представляющими разных пользователей.

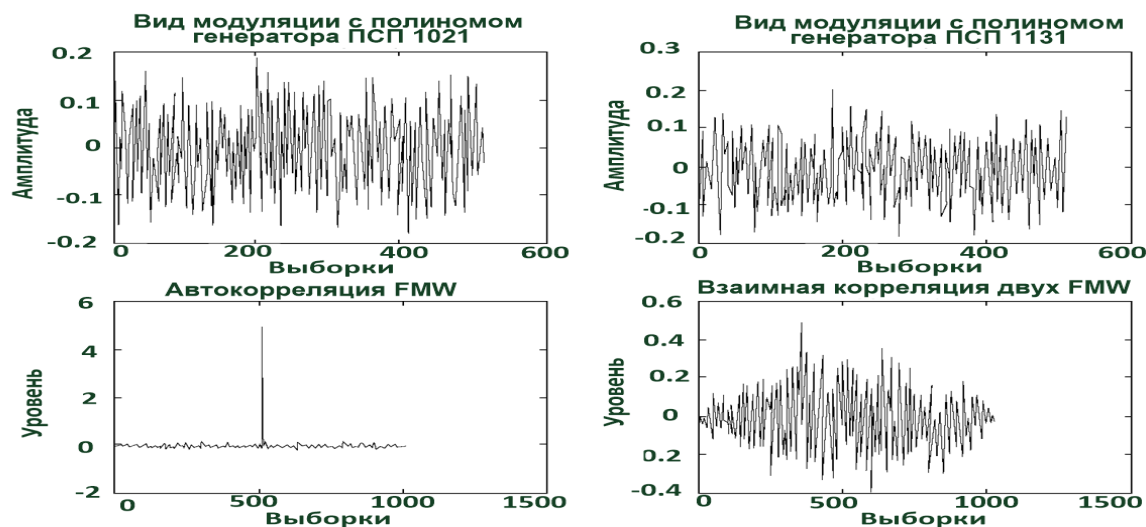
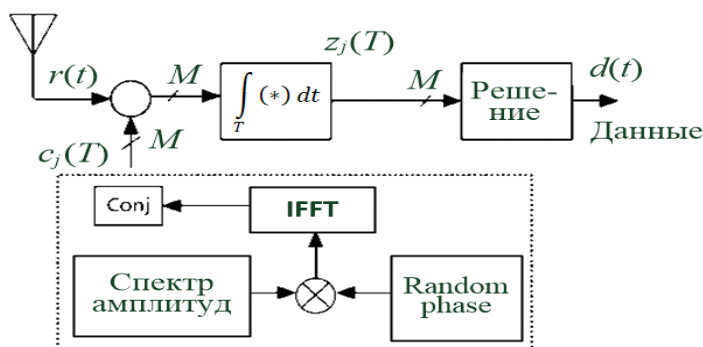


Рис. 3. Множественный доступ CDMA в системах TDCS

Автокорреляционная функция *FMW* является действительно импульсно-подобной, как ожидается от белого шума, и небольшой взаимокорреляционный отклик между пользователями двух разных ПСП является доказательством почти-ортогонального поведения, требуемого для множественного доступа [4]. Использование ПС кодирования в системе с *DS-CDMA* имеет целью расширение спектра, в то время как *TDCS* использует псевдослучайное кодирование для скремблирования сигнала с целью его крипто- и имитозащиты. В работе систем *TDCS* использовались два вида двоичной модуляции: противоположные сигналы (типа кодов Баркера) и ортогональная модуляция, именуемая манипуляцией циклическим сдвигом (*CSK - cyclic shift keying*) [4], которая имеет преимущества характеристик шумоподобных сигналов *FMW* (т.е., автокорреляционная функция *FMW* имеет вид как на рис. 3). Основанная на этом модуляция *CSK* в системах

TDCS использует циклические сдвиги формы волны *FMW* для представления разных символов. Для двоичной манипуляции (*BCSK*), первый символ, $S_1(t)$, является самым сигналом *FMW* и второй символ, $S_2(t)$, генерируется с помощью циклического сдвига сигнала *FMW* на одну половину его символьного периода, т.е. на $(T_s/2)$. Этот циклический сдвиг во временной области вызывает линейный сдвиг фазы в частотной области без влияния на амплитуду. В случае *M*-ичной манипуляции (*M-CSK*), первый символ $S_1(t)$ также является сигналом *FMW*, а j -ый символ $S_j(t)$ для $j = 1, 2, 3, \dots, M$ получают с помощью циклического сдвига сигнала *FMW* на $j \cdot T_s/M$, где M является объёмом алфавита сигнала *FMW*. *M*-ичная волна может быть рассмотрена как расширение двоичного случая [3].

Принимаемый сигнал $r(t)$ является суммой переданного сигнала $s(t)$, канального аддитивного белого гауссова шума (АБГШ) $n(t)$ и интерференции $i(t)$, если присутствует. Как и для любой системы связи, первым шагом обработки сигнала $r(t)$ в приёмнике *TDCS* перед демодуляцией является захват сигнала, обнаружение и синхронизация. Принятый сигнал $r(t)$ предварительно обрабатывается с использованием или прямой временной корреляции (*DTC*) или техники преобразования Фурье. Выход препроцессора затем передаётся к детектору для определения наличия сигнала и его выравнивания (коррекции). Если сигнал обнаружен и адекватно корректирован, то приёмник продолжает оценку сопровождения пакета и символьную синхронизацию. Структура демодулятора приёмника *TDCS* показана на рис. 4 с помощью пунктирной линии, отделяющей процессы генерации сигналов *FMW*, идентичных используемым в передатчике.



Местная *FMW*, M – число гипотез сигнала

Рис. 4. Функциональная схема приёмника *TDCS*

Сигнал $r(t)$ коррелируется с опорными сигналами местного гетеродина $c_j(T)$. Для двоичной модуляции ($M = 2$) существует один локально генерированный опорный сигнал для каждого возможного символа. Каждый из M корреляторов генерирует достаточную статистику - $z_j(T)$. Решения по критерию максимума правдоподобия (*ML*) используются для достаточной статистики для получения оценки данных - $d(t)$. Используемые правила принятия решения зависят от модуляции данных. Для противоположной модуляции, если $z(T)$ является положительным, то принимается $s_1(t)$; в противном случае принимается гипотеза $s_2(t)$. При манипуляции двоичным циклическим сдвигом (*BCSK*), если $z_1(2t) - z_2(t) > 0$, то принимается $s_1(t)$; в противном случае $s_2(t)$. Наконец, для *M*-ичной *CSK* принятый сигнал коррелируется с M опорными сигналами, и достаточная статистика с наибольшей амплитудой выбирается как принятый символ.

Технология *OFDM*, также основанная на быстром дискретном преобразовании Фурье (*FFT*) [5] является рабочей лошадкой сетей 4G и 5G высокоскоростной беспроводной широкополосной связи благодаря способностям по компенсации многолучевости с помощью простого защитного интервала в начале каждого символа без необходимости реализации эквалайзера. Сигналы с *OFDM* формируются в частотной

области и затем с помощью обратного быстрого преобразования Фурье (*IFFT* – Invers Fast Fourier Transform) переносятся во временную область для передачи по радиоканалу. Полученные на выходе модуля *IFFT* временные отсчёты через цифроаналоговый преобразователь (ЦАП) и фильтр передатчика поступают в непрерывный канал передачи как представлено на рис. 5.

Сигналы на поднесущих частотах ортогональны, поскольку разнос частот между ними равен $\Delta f = 1/T_U$, где $T_U = T_s - T_g$ – полезная длительность символа сигнала без учёта защитного интервала (T_g), называемого циклическим префиксом. На интервале T_U помещается целое число периодов колебаний любой k -й частоты поднесущей. В технологии OFDM один высокоскоростной поток данных заменяется множеством низкоскоростных медленных потоков с большими длительностями символов. Причём, чем уже полоса поднесущего колебания, тем больше длительность символа сигнала T_s и больше отношение сигнал/шум $h^2 = P_c T_s / N_0$. Сигнал OFDM, использующий большие длительности символа T_s имеет большую спектральную эффективность, благодаря использованию видов модуляции с большим числом гипотез сигнального созвездия M-QAM. Распределение частотно-временного ресурса сигналом OFDM представлено на рис. 5 [5].

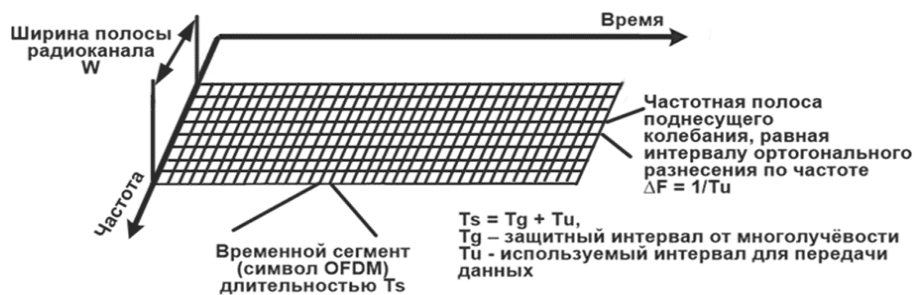


Рис. 5. Распределение частотно-временного ресурса сигналом OFDM

Блок схема генерации сигнала OFDM приведена на рис. 6 [7].

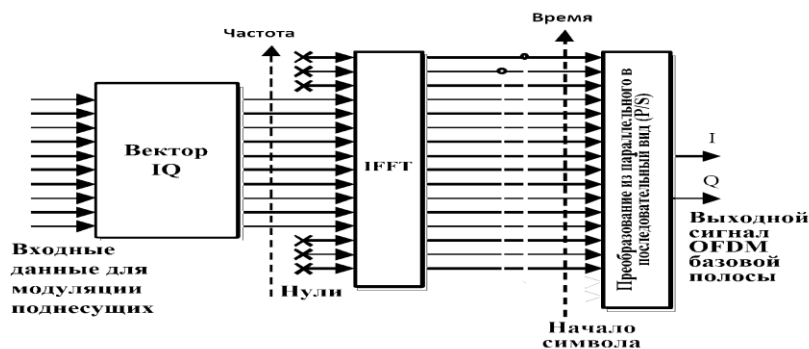


Рис. 6. Блок схема генерации сигнала OFDM

Комбинирование OFDM с мультиплексированием с помощью кодового разделения (CDM) нашло широкое использование в сотовых сетях связи в начале 2000-х [2], особенно схема модуляции MC-CDMA (Multi-Carrier-CDMA). Эта схема эффективно подавляет интерференцию многолучевости с помощью OFDM, на которой построена MC-CDMA, и в то же время, обеспечивает возможности множественного доступа с кодовым разделением. Поскольку все три технологии основаны на БПФ, но OFDM способна максимально приблизиться к границам спектральной эффективности К. Шеннона, и обеспечить максимальную скорость передачи за минимальное время слота, рассмотрим её подробнее. На рис. 7 приведена схема передатчика и приёмника OFDM.

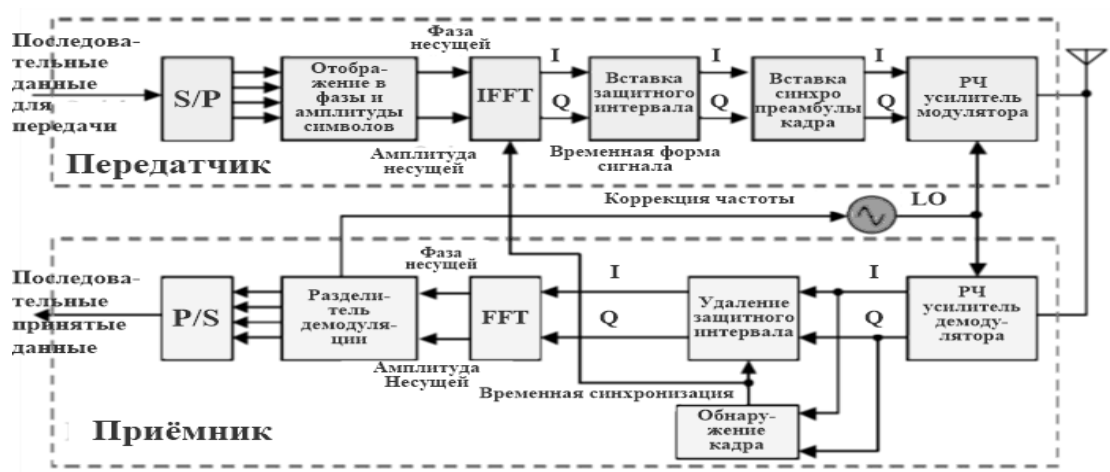


Рис. 7. Схема передатчика и приёмника OFDM

Выводы

Приведён краткий обзор технологий *OFDM* и *MC – CDMA*, также более детальное представление системы TDCS, архитектуры её передатчика и приёмника. Фундаментальное различие между этими тремя технологиями состоит в том, что технология *TDCS*, как динамическая технология частотно-адаптивной связи активно борется с помехами путём перемещения спектра передаваемого сигнала на чистые от помех рабочие частоты. Это повышает спектральную эффективность и помехозащищённость системы связи *TDCS*, поскольку для борьбы с помехами работают одновременно и приёмник и передатчик в радиолинии. Технологии когнитивного радио совместно с изменениями стратегий организации связи содержат много возможностей для решения проблемы частотной адаптации. Присущая *TDCS* способность очистки спектра для передачи данных и гибкость синтеза частотных доменов в системах *OFDM* и *MC – CDMA* делает эти три технологии идеальными кандидатами перспективного когнитивного радио *CR*. Поскольку все три системы работают в домене эффективного быстрого преобразования Фурье, располагают одинаковыми возможностями адаптации, которые необходимо использовать по обстановке в зависимости от требований приложений. Не следует забывать, что перегрузка спектра и разрывы линий связи вызываются в основном не отсутствием доступного спектра, а не способностью эффективно использовать то, что доступно. Поэтому целесообразно внедрить технологию частотной адаптации в системы *OFDM* и *MC – CDMA*, чтобы не передавать символы на поднесущих, поражённых помехами, например, как представлено на рис. 8, где 2-я, 3-я и 7-я поднесущие (сверху) (горизонтальные полосы), также 10-я, 11-я, 15-я не передаются, как занятые помехами. При этом количество поднесущих сохраняется, и они распределяются внутри более широкой полосы частот. Такие же рассуждения справедливы и для сигналов *MC – CDMA*. Построенные на базе *OFDM* они будут иметь такой же вид, как и сигнал *OFDM*, приведённый на рис. 8.

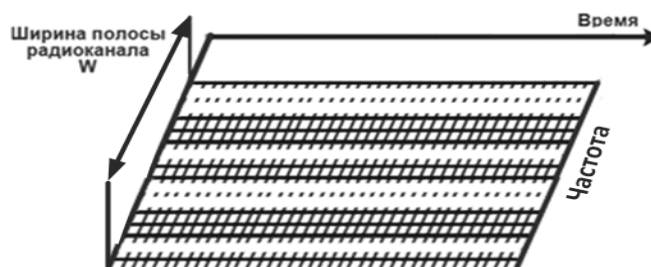


Рис. 8. Частотно-адаптивное распределение ресурсов сигналов *OFDM* и *MC – CDMA* при наличии помех в канале

Библиографический список

1. Mitola J. Cognitive radio architecture: the engineering foundations of radio XML // Wiley & Sons Inc. Publication, New Jersey. 2006. 488 p.
2. Hanzo L., M. Munster, B.J. Choi and T. Keller. OFDM and MC-CDMA for Broadband Multi-user Communications, WLANs and Broadcasting; Monograph. IEEE Press-John Wiley. 2002. 254 p.
3. Atarashi H., S. Abeta, M. Sawahashi, "Performance Evaluation of Coherent High-Speed TD-OFCDM Broadband Packet Wireless Access in Forward Link Employing Multi-Level Modulation and Hybrid ARQ", IEICE Trans. Fundamentals, vol. E84-A, no.7. July 2001. pp.1670-1680.,
4. Roberts M. L., M.A. Temple. R.A. Raines, E.P. Magee. Initial Acquisition Performance of a Transform Domain Communication System: Modeling and Simulation Result. MILCOM 2000, 21st Century Military Communications Conference Proceedings Volume: 2, Oct. 22–25, 2002, pp. 1119–23.
5. Рашич А.В. Сети беспроводного доступа WiMAX: учеб. пособие / Рашич А.В.— СПб.: Изд-во Политехн. ун-та, 2011. — 179 с.
6. Chakravarthy et al. "TDCS, OFDM, and MC-CDMA: A Brief Tutorial", in IEEE Communications Magazine, vol 43, Issue 9, Sept. 2005. pp. S11 -S16.
7. Adaptive Techniques for Multiuser OFDM. Thesis submitted by Eric Phillip LAWREY BE (Hons) in December 2001 for the degree of Doctor of Philosophy in Electrical and Computer Engineering School of Engineering James Cook University. 2001. p. 327.

Ya.S. Nekrasov, T.I. Goryacheva

Research of communication systems based on transformation domains

Nizhny Novgorod State Technical University n.a. R.E. Alekseev

The basic features of the structures of transmitters and receivers of cognitive radio systems that implement broadband modulation schemes based on the forward and inverse fast Fourier transforms (FFT) are considered. These include Orthogonal Frequency Division Multiplexing (OFDM) signals, Multi Carrier -code division multiple access (MC-CDMA) signals, and Fundamental Modulation Waveform (FMW) signals, used in Transform-Domain Communication System (TDCS). It is proposed to implement frequency adaptation algorithms to increase the spectral efficiency of communication systems with OFDM and MC-CDMA.

Keywords: orthogonal frequency division multiplexing, multi-carrier - code division multiple access, transform domain communication systems, fast Fourier transform, frequency adaptation.

П.Н. Кузьмичев², С.Ю. Меженин¹, Ю.С. Егоров¹, Е.О. Ульянова¹

ПОДХОД К ФОРМАЛИЗАЦИИ ПРОЦЕССА ПРОЕКТИРОВАНИЯ СЕТИ ТЕХНОЛОГИЧЕСКОЙ СВЯЗИ ДЛЯ ОБЪЕКТОВ ТРУБОПРОВОДНОГО ТРАНСПОРТА

¹ Нижегородский государственный технический университет им. Р.Е. Алексеева

² ООО «Газпром проектирование», г. Нижний Новгород

В работе рассмотрен подход к проектированию сети технологической связи для объектов трубопроводного транспорта, который позволит повысить эффективность принимаемых решений за счет оптимизации системно-сетевых и технико-экономических характеристик проектируемого объекта.

Ключевые слова: трубопроводный транспорт, сеть технологической связи, системно-сетевые решения, система поддержки принятия решений.

В настоящее время при проектировании сетей технологической связи трубопроводного транспорта часто возникают вопросы, связанные с эффективностью ее использования, что в ряде случаев ставит под сомнение целесообразность строительства того или иного объекта. На практике возникает необходимость оптимизации системно-сетевых и технико-экономических характеристик проектируемого объекта, определяющих его функционирование. Выбор характеристик не может производиться изолированно, а напротив должен осуществляться комплексно с учетом работы проектируемого объекта, как составного элемента большой распределенной сети связи [1-2].

С целью обеспечения организационно-технического единства и целостности систем связи трубопроводного транспорта на начальном этапе проектирования осуществляется разработка системно-сетевых решений по организации связи. Такие работы охватывают большую область знаний, включая сетевое планирование, технологии передачи информации, типы используемого оборудования, выбор среды передачи информации [3]. Требуется проведение анализа большого объема данных в части текущего состояния сети, выполненных и планируемых к выполнению проектов, планов развития сети связи на ближайшую и отдаленную перспективу, тенденций в использовании новых технологий передачи информации, технико-коммерческих предложений поставщиков оборудования.

При проектировании сетей технологической связи на выбор проектных решений в соответствии с требованиями технического задания (далее – ТЗ) определяющее влияние оказывает тип проектируемых линий связи (РЛЛ – радиорелейная линия связи, ВОЛС – волоконно-оптическая линия связи, спутниковая линия связи и т.д.) [4]. При этом выбор осуществляется исходя из типовых системных решений по применению данных линий и их сравнения. В рамках развиваемого подхода предлагается каждый новый проект сравнивать с шаблонами в «архиве проектов» по заранее определенным (типовым) характеристикам.

Установление степени соответствия нового проекта одному или нескольким паттернам из архива проектов, предопределяет возможность использования опыта существующих наработок и позволит значительно снизить затраты ресурсов – временных, финансовых, а также человеческих. Кроме того, в значительной мере уменьшаются проектные риски за счет снижения вероятности, что будут допущены несоответствия или ошибки на различных этапах проектирования, что в свою очередь, позволит снизить стоимость доработок и исключить повторное перепроектирование. Известно, что чем раньше были допущены несоответствия, тем выше стоимость их исправления в будущем.

В результате специалистам предприятия будет доступна владеть необходимая база актуальных знаний, включая технологии и шаблоны проектирования.

Для выбора прототипа из множества ранее выполненных проектов необходимо определить критерий и сформировать постановку задачи дискретной оптимизации. Поскольку задача является многокритериальной, то первым этапом решения может быть определение множества альтернатив, оптимальных в смысле Парето (множество Парето) $A_{\text{П}}$. Выделение области допустимых решений $A_{\text{Д}}$, удовлетворяющих всем ограничениям, позволяет отбросить заведомо неудовлетворительные (недопустимые) $A_{\text{Н}}$ альтернативы $A_{\text{Д}} \cup A_{\text{Н}} = A$. Принятые ранее проектные решения будем считать допустимыми, если обеспечено соответствие требованиям ТЗ на проектирование нового (планируемого) объекта, что можно представить в виде условий:

$$Q \geq Q_{\text{ТЗ}}, \quad T \leq T_{\text{ТЗ}}, \quad C \leq C_{\text{ТЗ}}, \quad (1)$$

где Q – обобщенный показатель технического качества проекта, связанный с функциональными требованиями, T – срок разработки, C – стоимость проектирования; $Q_{\text{ТЗ}}$ – требуемое значение обобщенного показателя функциональных требований ТЗ на новый проект, $T_{\text{ТЗ}}$ – заданная продолжительность срока разработки нового проекта, $C_{\text{ТЗ}}$ – приемлемые затраты на проектирование нового объекта.

Условие (1) можно использовать для определения множества допустимых решений

$$A_{\text{Д}} = \{a \in A | Q(a) \geq Q_{\text{ТЗ}}, T(a) \leq T_{\text{ТЗ}}, C(a) \leq C_{\text{ТЗ}}\}. \quad (2)$$

Таким образом, все альтернативы $a \in A_{\text{Д}}$ удовлетворяют требованиям ТЗ.

Если есть альтернативы (проекты), для которых требования по всем характеристикам

выполняются, то $Q = 1$ и проекты сравниваются по T и C .

Трудоемкость T и стоимость C проектирования нового объекта складывается, соответственно, из затрат времени и стоимости разработки системно-технических решений (СТР) и проектной документации (ПД):

$$T = T_{\text{СТР}} + T_{\text{ПД}}; \quad C = C_{\text{СТР}} + C_{\text{ПД}}.$$

Если есть альтернативы (проекты), для которых все требования выполняются, то $Q = 1$, этап разработки системно-технических решений практически исключается и возможно сокращения сроков и стоимости проектирования на величины $\Delta T \approx T_{\text{СТР}}$ и $\Delta C \approx C_{\text{СТР}}$. Если отсутствуют альтернативы (проекты), для которых все требования выполняются и $Q < 1$, то разработка системно-технических решений выполняется с частичным переиспользованием ранее принятых решений и $0 < \Delta T < T_{\text{СТР}}$, $0 < \Delta C < C_{\text{СТР}}$. Величины ΔT и ΔC характеризуют выигрыш от предложенной процедуры интеллектуальной поддержки системно-технических решений.

В этом случае при определении проекта-аналога для последующей модернизации используется подход и алгоритм, разработанный в [5] и развитый в [6], основанный на применении нечеткой логики и функций принадлежности специального вида.

Для решения задачи многокритериального выбора проекта аналога определим множество альтернатив, оптимальных в смысле Парето (множество Парето) $A_{\text{П}}$ и будем искать решение среди Парето-оптимальных альтернатив, удовлетворяющих требованиям ТЗ (2), т.е. на пересечении множеств $A_{\text{П}} \cap A_{\text{Д}}$. Если множество $A_{\text{П}} \cap A_{\text{Д}}$ содержит несколько альтернативных проектов, то их ранжирование будет зависеть от способа решения задачи многокритериального выбора.

Так выбор наилучшего варианта может осуществляться путем сведения многокритериальной задачи выбора к однокритериальной с использованием дополнительной информации, отражающей предпочтения проектировщика – лица, принимающего решения (ЛПР). Перейти к однокритериальной задаче позволяют методы

свертки векторного критерия и методы ранжирования критериев. Достаточно широко применяется метод аддитивной свертки критериев, основанный на оптимизации (например, максимизации) взвешенной суммы нормированных частных целевых функций

$$\bar{J}(a) = \sum_{i=1}^3 w_i J_i(a), \quad (3)$$

где $J_i(a)$ – нормированные частные целевые функции, w_i – веса, характеризующие

важность частных целевых функций такие, что $\sum_{i=1}^3 w_i = 1$.

Нормировку частных целевых функций предлагается выполнить так, чтобы максимизация (3) обеспечивала максимизацию относительного «запаса» выполнения ограничений (1) с учетом важности отдельных частных целевых функций

$$J_1(a) = (T_{ТЗ} - T(a))/T_{ТЗ}, J_2(a) = (C_{ТЗ} - C(a))/C_{ТЗ}, J_3(a) = (Q(a) - Q_{ТЗ})/Q_{ТЗ}. \quad (4)$$

Другой, более простой подход к решению задачи многокритериального выбора основан на определении проектировщиком главной целевой функции. Так при выборе в качестве главной целевой функции $J_{Г}(a) = J_2(a)$ задачу поиска наилучшего проекта-аналога можно представить в виде

$$\hat{a} = \arg \max_{a \in A_{П} \cap A_{Д}} J_2(a) = \arg \min_{a \in A_{П} \cap A_{Д}} C(a). \quad (5)$$

Основной ожидаемый эффект в снижении проектного риска $R = P_{ош} C_{Иош}$ и средней стоимости проектирования за счет переиспользования системно-технических решений, и минимизации «переделок», а также в увеличении доли проектов, выполненных в срок.

Таким образом, на основе предлагаемого подхода при проектировании нового объекта обеспечивается возможность повторно использовать принятые ранее системно-технические решения, соответствующие требованиям нового ТЗ. Интеллектуальная поддержка процесса проектирования сети технологической связи для объектов трубопроводного транспорта, будет способствовать повышению эффективности принимаемых решений за счет использования корпоративной базы знаний, содержащей сведения о ранее выполненных проектах («архив проектов») и онтологическую модель процесса проектирования, позволяющую применять знания, формализованные в виде правил принятия решения.

Библиографический список

1. Горев П.В., Кузьмичев П.Н., Меженин С.Ю. Определение сетевых требований к объектам связи // XVIII Международный научно-промышленный форум «Великие реки – 2016»: Непрерывное профессиональное образование в сфере устойчивого развития», 17 – 20 мая 2016 г. / ННГАСУ; отв. ред. А.А. Лапшин. – Н. Новгород: ННГАСУ, 2016. С. 199-202.
2. Савченков С.В., Смычѐк М.А., Милов В.Р., Новиков Н.В. Способы оценки и анализа радиопокрытия проектируемой сети технологической подвижной радиосвязи // Электросвязь. 2011. № 2. С. 30-33.
3. Ремешков В.Ю., Милов В.Р. Метод многокритериального выбора оптимального проектного решения по технологической связи магистральных газопроводов с учетом надежности направлений информационного обмена // Электронный научный журнал Нефтегазовое дело. 2015. № 4. С. 132-150.
4. Ремешков В.Ю., Писарев О.В., Милов В.Р. Применение автоматизированного алгоритма расчета коэффициентов готовности двухполосных сетей для анализа надежности технологических сетей связи в газовой отрасли // Современные проблемы науки и образования. 2014. № 3. С. 152.

5. Баранов В.Г., Седаков Д.В., Милов В.Р. Способ управления конфигурацией информационных систем // Информационно-измерительные и управляющие системы. 2017. № 8 (15). С. 26-33.

6. Скрипник И.В., Егоров Ю.С. Выбор системно-технических решений при построении радиосвязного оборудования с применением аппарата нечеткой логики // Проектирование и технология электронных средств. 2019. № 3. С. 21-27.

P.N. Kuzmichev, S.Yu. Mezhenin, Yu.S. Yegorov, E.O. Ulyanova

**APPROACH TO THE TECHNOLOGICAL COMMUNICATION NETWORK DESIGN
FORMALIZATION FOR PIPELINE TRANSPORT OBJECTS**

*Nizhny Novgorod state technical university n.a. R.E. Alekseev
OOO Gazprom Proektirovaniye, Russia, Nizhny Novgorod*

This paper proposes methods for an intellectual support of decision-making management. The method allows increasing the efficiency of decisions and optimizing the structural and technical characteristics of the pipeline transport object.

Keywords: pipeline transport, technological communication network, system-network solutions, decision support system.

МЕТОДЫ ВЗАИМОДЕЙСТВИЯ ЭЛЕМЕНТОВ ПОРТАЛЬНЫХ СТАНКОВ, ОСНАЩЕННЫХ ЧИСЛОВЫМ ПРОГРАММНЫМ УПРАВЛЕНИЕМ

Нижегородский государственный технический университет им. Р.Е. Алексеева

Рассмотрена структура станочного комплекса портального типа, методы взаимодействия его компонентов. Оценена возможность модернизации компонентов станочных комплексов.

Ключевые слова: станок, ЧПУ, координатный стол, рабочий инструмент, система управления.

Введение

В настоящее время станки – это необходимость, поскольку новый виток развития технологического уклада требует производить больше за те же деньги. Человек работает 8 часов в день, станок – 24. Человеку нужно время для переобучения, станок меняет программы в кратчайшие сроки. По большей части, в России станки на малых и средних предприятиях старые, но механика до сих пор в хорошем состоянии, а именно она, по сути, представляет собой основную часть стоимости станка. При этом на новых станках механика существенно не отличается – применяются типовые решения. Электроника, однако, неизбежно устаревает, и проблема модернизации станков с ЧПУ заключается в применении современных систем управления.

Рассмотрим типовую комплексную систему станка портального типа. Она включает в себя три обязательных составляющих: координатный стол, рабочий инструмент и управляющую систему.

Координатный стол

Координатный стол включает в себя механические и электрические компоненты. К механическим компонентам относятся передача и линейные направляющие, вместе образующие систему линейного перемещения. В качестве линейных направляющих, как правило, используются профильные рельсы, по ним движется портал станка. Передачи обычно ставятся реечные зубчатые, но также могут использоваться ременные или винтовые. Координатный стол портального стола чаще всего имеет три оси: X, Y, Z; а для их перемещения применяются 4 привода: 2 двигают портал по оси X, один привод управляет осью Y, и еще один осью Z [1].

К электрическим компонентам координатного стола относятся драйверы управления приводами и концевые датчики, ограничивающие максимальный и минимальный ход каждой оси. Драйверы приводов осей современных станков получают один сигнал движения в качестве входного и сигнал направления движения. Наиболее распространен промышленный протокол STEP/DIR. В этом случае не имеет значения, какой привод установлен на той или иной раме. Это может быть привод шагового двигателя с обратной связью энкодера или без нее. Это также может быть сервопривод. Выбор приводного решения повлияет на характеристики механических частей станка, а также на точность, скорость, ускорение и, следовательно, на шум, возникающий при движении оси станка [2].

Как правило, для взаимодействия с одной осью станка применяются следующие сигналы: STEP, DIR, ENABLE, MIN, MAX, HOME [1,2].

На рис. 1 представлена схема входных и выходных сигналов для одной оси станка (ось X).

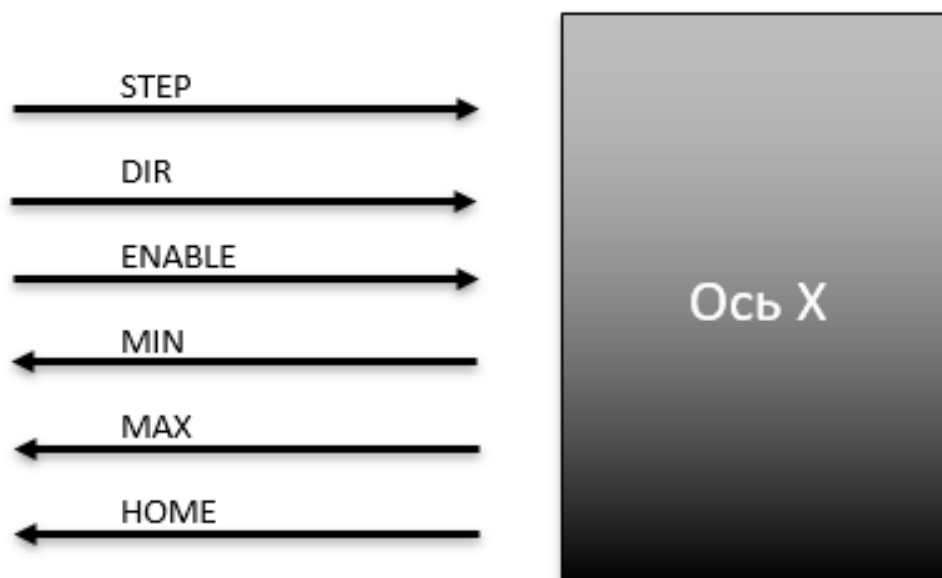


Рис. 1. Типовая схема входных и выходных сигналов для станочной оси

Сигналы STEP, DIR и ENABLE подаются приводом оси, взаимодействующим с мотором двигателя. Сигнал ENABLE опционален, если он подаётся, то снимается напряжение с выхода драйвера и двигатель останавливается. Если этот функционал не требуется, то вход драйвера просто не подключают [3]. Сигналы MIN, MAX — это сигналы концевых датчиков, указывающих минимальное и максимальное аварийные положения машины, и датчика HOME, расположенного между ними.

На рис. 2 представлена схема расположения концевых датчиков для одной оси станка.



Рис.2. Типовая схема расположения датчиков MIN, MAX, HOME для одной оси станка

Обычно датчик HOME отражает координату 0 оси, а программный лимит перемещения оси рассчитывается на основе этой координаты. В некоторых случаях MIN и HOME физически являются одним датчиком. Это решение является неудачным, потому что, если система управления зависает во время выполнения процесса поиска нуля, автоматическая защита от выхода за лимит будет отключена, и неизбежно произойдет столкновение с физическим ограничителем – ножкой станка. Эта ситуация потенциально опасна и может привести к изгибу рамы или ограничителя из-за столкновения механических частей [4].

В настоящее время контроллер осей без проблем умещается в одну программируемую плату, причем затраты в таком случае идут только на саму плату, так как в открытом доступе уже есть готовые прошивки с открытым исходным кодом, такие как grbl, g2core, Marlin, tinyG. Управлять станком в таком случае можно прямо с компьютера через интерфейс USART [1].

При модернизации станков может возникнуть проблема в том, что приводы поддерживают только аналоговое управление, в то время как контроллер осей имеет возможность подавать лишь STEPDIR сигналы. В таком случае необходимо на каждый аналоговый привод ставить ПИД-регулятор.

Из существующих реализаций таких регуляторов можно выделить DC Servomotor Controller by Elm-Chani YAPSC:10V. Они превращают STEPDIR-импульсы в аналоговый

сигнал $\pm 10V$, позволяя управлять аналоговыми приводами. Оба решения имеют открытый исходный код, позволяя модифицировать прошивку под свои нужды, но YAPSC:10V является более современным проектом.

Рабочий инструмент

Рабочий инструмент определяет принадлежность машины к виду работ. Наиболее часто применяются следующие инструменты: фрезер, плазморез, лазер, нож и т.п.

У любого рабочего инструмента есть минимум один канал управления – разрешение движения. Если присутствует дополнительный интерфейс для тонкой настройки инструмента, то его производитель поставляет программное обеспечение.

На рис. 3 представлена схема взаимодействия системы управления с рабочим инструментом.



Рис. 3. Типовая схема взаимодействия системы управления с рабочим инструментом

Поскольку каждый рабочий инструмент имеет свой собственный набор задач, интерфейс для настройки каждого типа рабочего инструмента различается от производителя и модели инструмента [5].

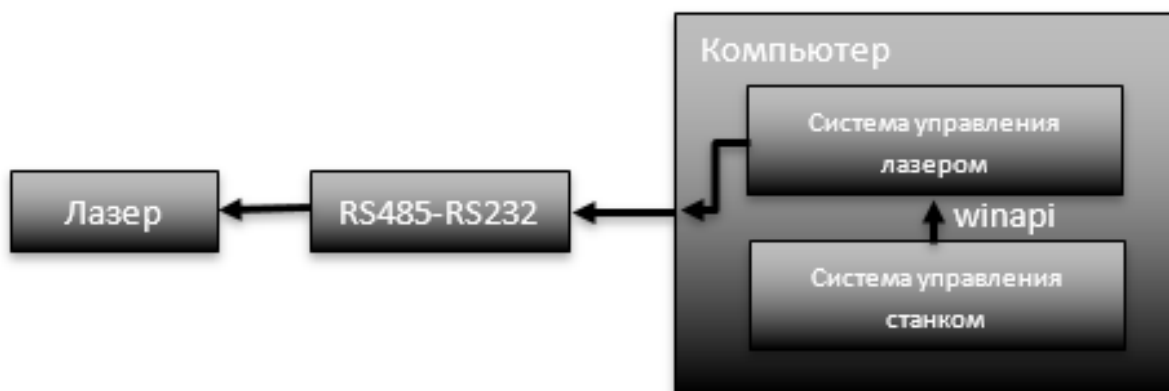


Рис. 4. Схема взаимодействия системы управления с лазером TL800

Например, газоразрядный лазер TL800 переключается в режим импульса лишь щелчком по кнопке в поставленной производителем программе, аппаратно это сделать невозможно. Тогда любая система управления станком с ЧПУ не сможет управлять лазером, если только у неё нет возможности использовать winapi для управления окном ПО лазера.

На рис. 4 представлена в качестве примера схема взаимодействия системы управления с лазером TL800.

Система управления

Система управления взаимодействует с координатным столом и рабочим инструментом и выполняет команды, заранее подготовленные и сохраненные в виде файлов. Общепринятым стандартом взаимодействия с системой управления ЧПУ является G-code [2].

G-code подразделяются на 2 вида: G относящиеся к координатному столу и M относящиеся к рабочему инструменту. Команды M отличаются для разных типов станков в зависимости от рабочего инструмента [1].

Существует множество бесплатных и коммерческих программ, позволяющих посылать G-codeкоманды с компьютера на оборудование.

На рис. 5 представлена схема взаимодействия компонент станочного комплекса.

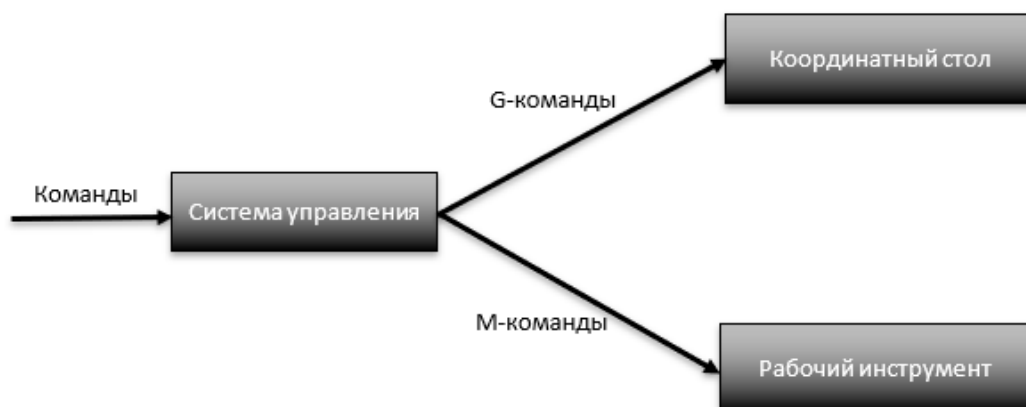


Рис. 5. Схема взаимодействия компонент станочного комплекса

Однако, роль системы управления станком с ЧПУ не обязательно должна сводиться лишь к обработке и отправке программ на станок. К ней может быть добавлен дополнительный функционал, оптимизирующий процесс производства в целом. Например, можно оборудовать станок видеокамерами и использовать систему управления для обработки изображения с этих камер. Полученные данные можно использовать для автоматизации процесса позиционирования координатного стола станка с ЧПУ в точке нуля заготовки, или для предотвращения ситуаций, когда головка лазера задевает за прогнувшийся металл и повреждается кристалл, что приводит к 40-минутному простоя на смену кристалла. Также можно добавить функционал, позволяющий системе управления самой составлять оптимальное расписание выполнения загруженных в неё программ обработки. Другими словами, систему управления станком с ЧПУ можно вполне превратить в интеллектуальную систему управления производством.

Выводы

С целью получения качественных параметров при модернизации порталных станков с ЧПУ необходимо:

- использовать драйверы управления приводами, поддерживающие протокол STEP/DIR;
- к приводам, поддерживающим управление только аналоговыми сигналами, подключать ПИД-регуляторы, используя открытые реализации DC Servomotor Controller или YAPSC:10V;
- не совмещать датчики MINи HOMЕв один;
- если у рабочего инструмента имеются каналы управления кроме разрешения движения, запрашивать у производителя рабочего инструмента документацию по каналам связи с инструментом;
- использовать G-codeдля описания программ, посылаемых на станок;
- использовать системы управления с дополнительным функционалом, оптимизирующим процесс работы станка или производства в целом.

Библиографический список

1. Устройство числового программного управления NC-201: Руководство по эксплуатации. – СПб., 2005. – 80 с.: ил.
2. Волчкевич Л.И. Автоматизация производственных процессов / Л. И. Волчкевич. - учеб. пособие. - М.: Машиностроение, 2005. - 380 с.
3. Пестов С. П. Проектирование технологических процессов обработки деталей на станках с ЧПУ / С. П. Пестов - Челябинск, 2002. – 66 с.
4. Бжозовский Б.М. Управление станочными комплексами/ Б.М. Бжозовский, В.В Мартынов. Саратов: Из-во Саратовского гос. техн. ун-та, 2004. –44 с.
5. Базров Б.М. Модульная технология в машиностроении / Б.М. Базров. – М.: Машиностроение, 2001. 368 с.

N.A. Metelkov, N.E. Posobilov

METHODS OF INTERACTION WITH COMPONENTS OF PORTAL TYPE COMPUTER NUMERIC CONTROL MACHINES

Nizhny Novgorod state technical university n.a. R.E. Alekseev

The paper describes the structure of typical portal type computer numeric control machine and the methods of interacting with its' components, points out the possibilities of modernization of machines.

Keywords: machine tool, CNC, coordinate table, machine tool, control system.

СРАВНИТЕЛЬНЫЙ АНАЛИЗ ЭФФЕКТИВНОСТИ АЛГОРИТМОВ МАШИННОГО ОБУЧЕНИЯ

Нижегородский государственный технический университет им. Р.Е. Алексеева

Работа посвящена сравнению эффективности алгоритмов машинного обучения. Рассмотрены несколько алгоритмов машинного обучения: полиномиальная регрессия; дерево принятия решений; случайный лес. Рассмотрены принципы их работы на базе данных поведения городского транспорта.

Ключевые слова: машинное обучение, дерево принятия решений, полиномиальная регрессия, случайный лес.

Введение

Машинное обучение — процесс, в результате которого машина (компьютер) способна показывать поведение, которое в нее не было явно заложено (запрограммировано) [1]. Главной задачей машинного обучения является автоматическое определение взаимозависимостей между наблюдаемыми и скрытыми переменными объекта, с тем, чтобы для произвольного объекта по его наблюдаемым компонентам можно было оценить возможные значения скрытых компонент. Цель машинного обучения – научить машину (точнее, программу) решать задачу, предъявив ей несколько примеров (с правильными и неправильными решениями). В машинном обучении применяются различные алгоритмы, причём определённые алгоритмы считаются достаточно многофункциональными. На эффективность алгоритмов влияет множество факторов: размер; структура набора данных. В данной работе рассмотрены алгоритмы машинного обучения: полиномиальная регрессия; дерево принятия решений; случайный лес, их эффективность для различных классов задач.

Полиномиальная регрессия

Полиномиальная регрессия может применяться в математической статистике при моделировании трендовых составляющих временных рядов. Временной ряд — это, по сути, ряд чисел, которые зависят от времени. Например, средние значения температуры воздуха по дням за прошедший год, или доход предприятия по месяцам. Порядок моделируемого полинома оценивается специальными методами, например, критерием серий. Цель построения модели полиномиальной регрессии в области временных рядов всё та же – прогнозирование.

Рассмотрим задачу полиномиальной регрессии в общем виде (рис.1). Все рассуждения основаны на обобщении рассуждений в задачах линейной и параболической регрессии.

Дерево принятия решений

В алгоритме «Дерево принятия решений» каждый лист представляет собой значение целевой переменной, изменённой в ходе движения от корня по листу. А каждый внутренний узел соответствует одной из входных переменных. Дерево может быть также «изучено» разграничением исходных наборов переменных на подмножества, которые основаны на тестировании значений атрибутов. Данный процесс повторяется на каждом из полученных подмножеств. Рекурсия завершается тогда, когда подмножество в узле имеет те же значения целевой переменной, таким образом, оно не добавляет ценности для предсказаний. Процесс, идущий «сверху вниз», индукция деревьев решений (TDIDT), является примером поглощающего «жадного» алгоритма, и на сегодняшний день является наиболее распространённой стратегией деревьев решений для данных, но это не

единственная возможная стратегия. В интеллектуальном анализе данных, деревья решений могут быть использованы в качестве математических и вычислительных методов, чтобы помочь описать, классифицировать и обобщить набор данных. Простым примером работы алгоритма служит решение о принятии предложения работы (рис. 2).

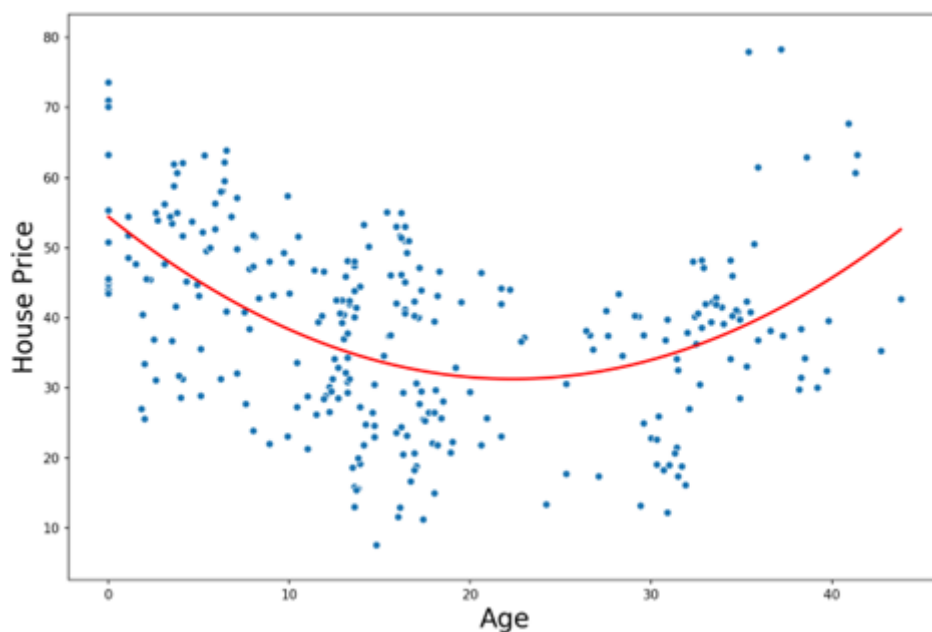


Рис. 1. Результаты реализации алгоритма машинного обучения «Полиномиальная регрессия»

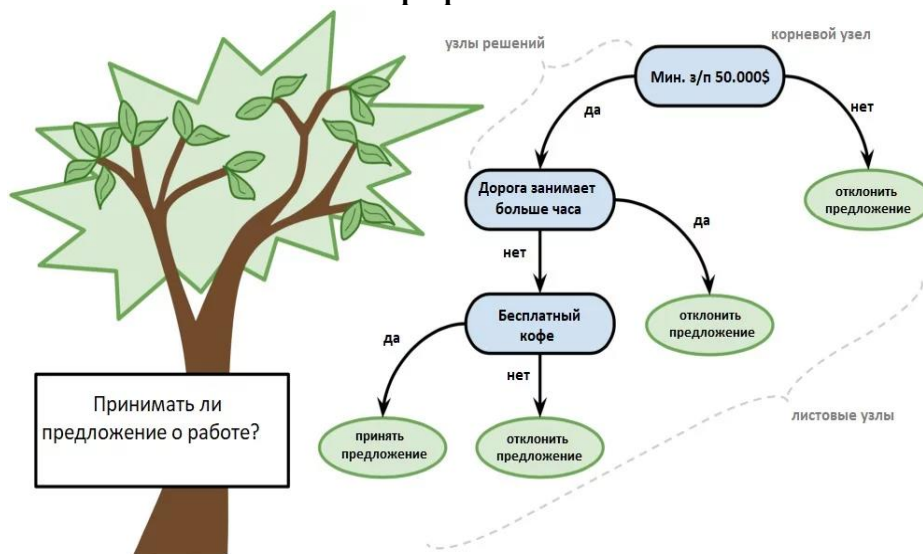


Рис. 2. Пример функционирования алгоритма машинного обучения «Дерево принятия решений»

Случайный лес

Случайный лес — это один из самых потрясающих алгоритмов машинного обучения, придуманные Лео Брейманом и Адель Катлер ещё в прошлом веке. Он дошёл до нас в «первозданном виде» (никакие эвристики не смогли его существенно улучшить) и является одним из немногих универсальных алгоритмов. Универсальность заключается, во-первых, в том, что он хорош во многих задачах, во-вторых, в том, что есть случайные леса для решения задач классификации, регрессии, кластеризации, поиска аномалий, селекции признаков и т.д.

Случайный лес — это множество решающих деревьев. В задаче регрессии их ответы усредняются, в задаче классификации принимается решение голосованием по большинству.

Случайный лес сочетает сотни или тысячи деревьев принятия решений, обучая каждое на отдельной выборке данных, разделяя узлы в каждом дереве с использованием ограниченного набора параметров. Итоговый прогноз делается путём усреднения прогнозов от всех деревьев (рис. 3).

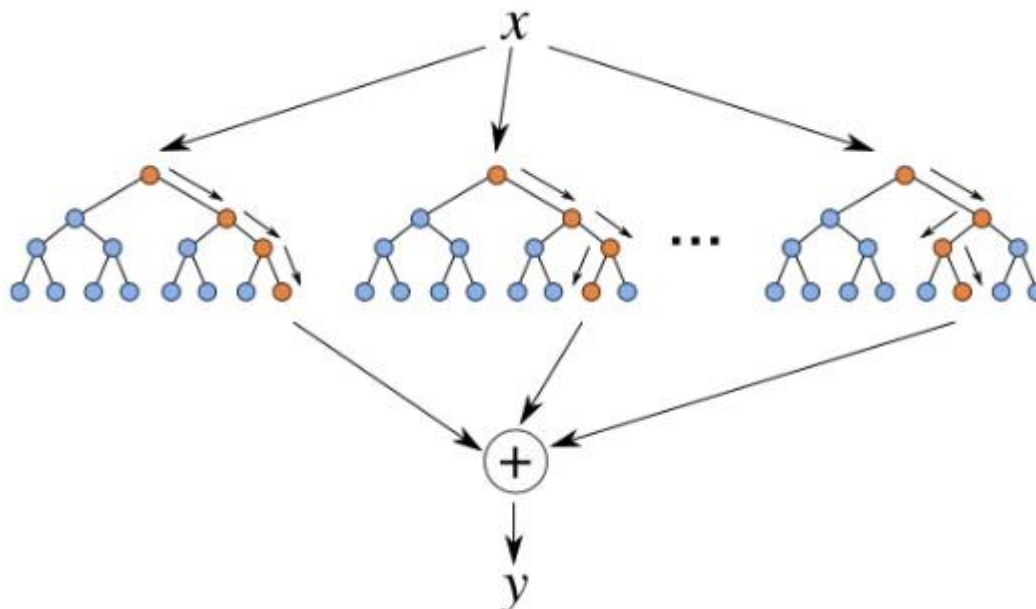


Рис. 3. Общий вид алгоритма принятия решений «Случайный лес»

О наборе данных

База данных содержит записи поведения городского движения города Сан-Паулу в Бразилии с 14 декабря 2009 года по 18 декабря 2009 года (с понедельника по пятницу). Регистрируется с 7:00 до 20:00 каждые 30 минут. «Час» можно разделить на утро, полдень, вечер, ночь, которые могут быть дополнительными кодами, использующими однократное кодирование. Дадим нумерацию событий, произошедших на дороге: 1. Сломанный грузовик; 3. Обездвиженный автобус; 4. Избыток автомобилей; 5. Жертва аварии; 9. Перебегающий пешеход; 11. Пожарные машины; 12. Происшествие с грузом; 14. Инцидент с опасным грузом; 16. Отсутствие электричества; 17. Огонь; 18. Точка затопления; 19. Проявления; 20. Неисправность в сети троллейбусов; 21. Дерево на дороге; 25. Семафор выключен; 27. Прерывистый семафор.

Отсутствие некоторых номеров связано с тем, что отсутствующие события не замедляют поток трафика.

Результаты моделирования

На графиках (рисунки 4-6) по оси X-происшествия на дороге, ось Y-процент замедления движения. Оранжевые столбцы – прогнозируемые, синие – действительные. Соотношение прогнозируемых данных к действительным должно быть 1:1. Таким образом, какой метод покажет лучшее соотношение, тот и будет определен лучшим по точности по сравнению с остальными методами.

Результаты прогнозирования различными методами машинного обучения представлены на рисунках 4-6.

Построение модели методом полиномиальной регрессии (рис.4).

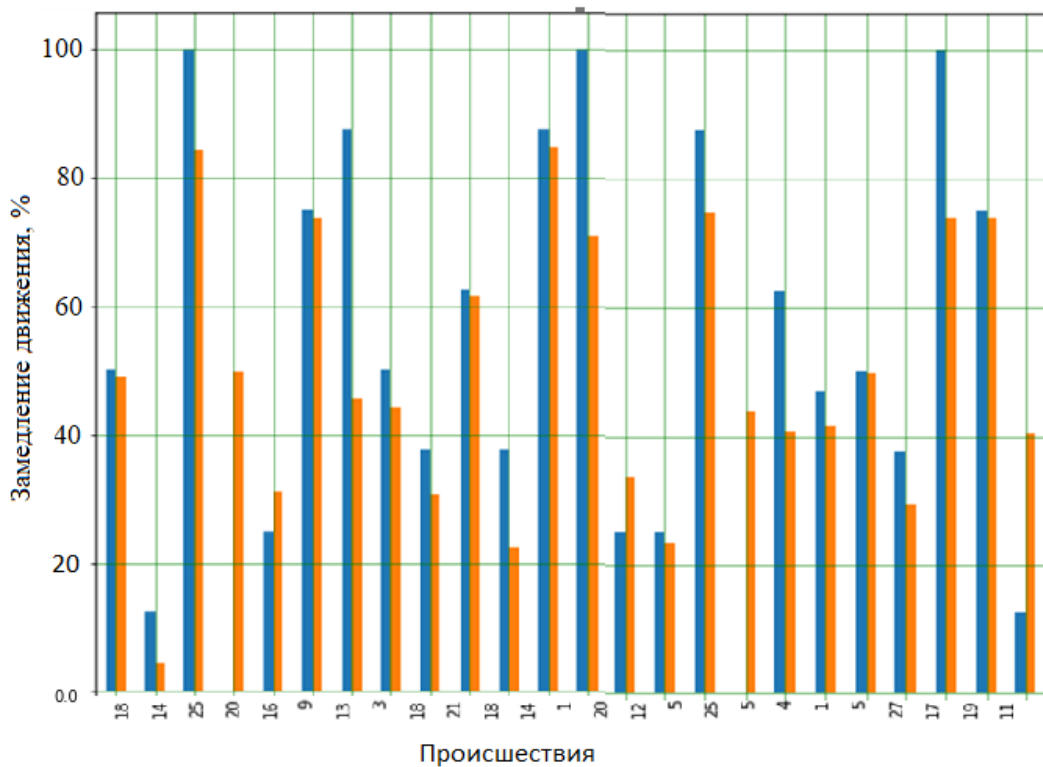


Рис. 4. Результат моделирования алгоритма принятия решений «Полиномиальная регрессия»

Построение модели при помощи дерева принятия решений (рис.5).

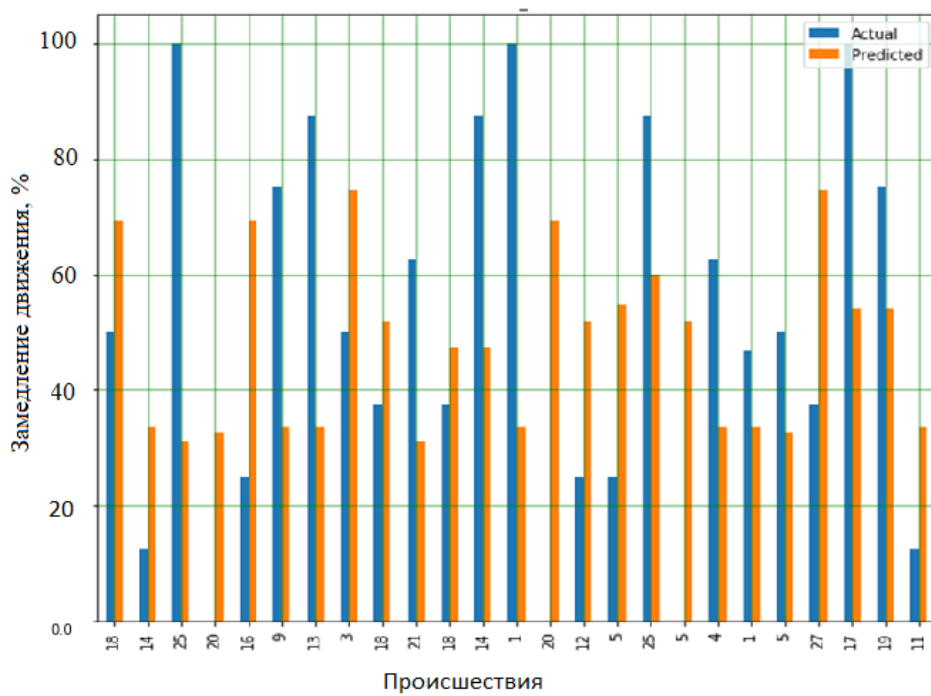


Рис. 5. Результат моделирования алгоритма принятия решений «Дерево принятия решений»

Построение модели при помощи метода случайного леса(рис.6)

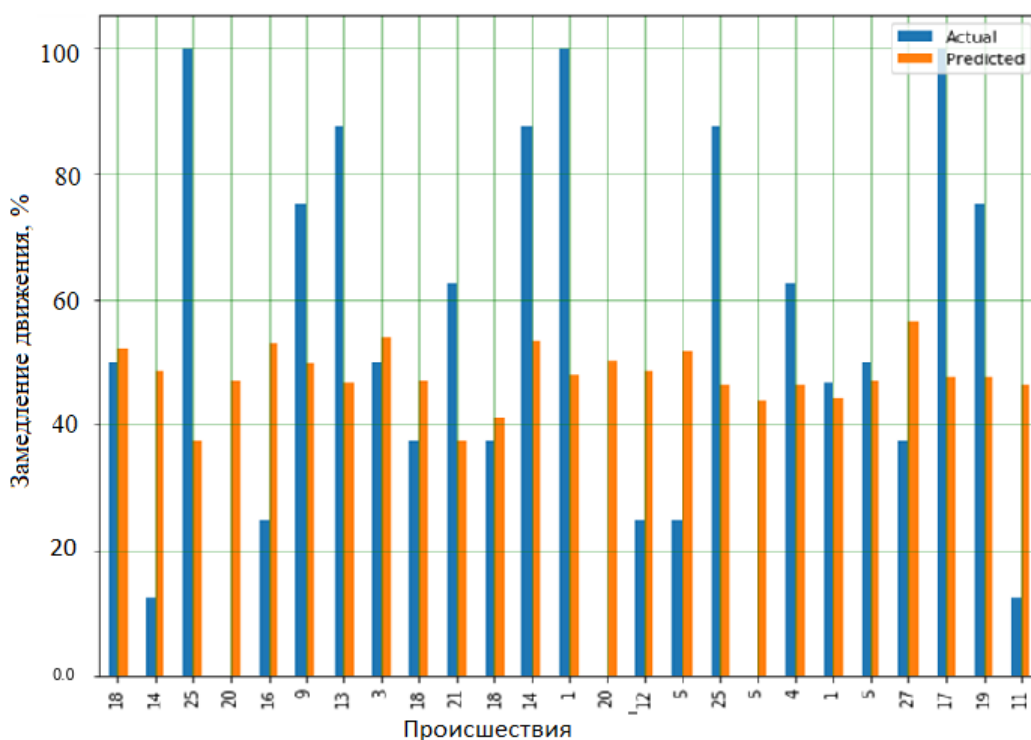


Рис.6. Результат моделирования алгоритма принятия решений «Случайный лес»

Исходя из результатов, наиболее эффективным методом является полиномиальная регрессия. Этот метод наиболее точно спрогнозировал поведение городского транспорта. Соотношение столбцов прогнозируемых данных (оранжевые) к действительным (синие) больше приближено к правильному (1:1) на графике метода полиномиальной регрессии (рис. 4).

Выводы

В работе были рассмотрены алгоритмы машинного обучения: полиномиальная регрессия, дерево решений, случайный лес. Также были рассмотрена их точность прогноза на примере базы данных поведения городского транспорта. Лучшую точность прогноза показал метод полиномиальной регрессии, именно он показал лучшее соотношение прогнозируемых данных к действительным.

Библиографический список

1. Chang S. What is the machine learning /S. Chang, T. Cohen, B. Ostdiek - Physical Review D. 2018 T. 97 № 5 С. 054024.
2. Kumar N.K. Ocean wave height prediction using ensemble of extreme learning machine/ N. K. Kumar, A. Mamun, R. Savitha - Neurocomputing. 2018. Т. 277. С. 12-20.
3. Китов В. Практические аспекты машинного обучения / Китов В.// Открытые системы. СУБД. 2016. № 1. С. 14-17.
4. Жуков Д.А. Задачи обеспечения эффективности машинного обучения при диагностике технических объектов/ Д.А. Жуков, В.Н. Клячкин - Современные проблемы проектирования, производства и эксплуатации радиотехнических систем. 2016. № 10. С. 172-174.
5. Голощапова И.О. Оценка инфляционных ожиданий российского населения методами машинного обучения / И.О. Голощапова., М.Л. Андреев - Вопросы экономики. 2017. № 6. С. 71-93.

K.S. Romanov, N.E. Posobilov

COMPARISON OF EFFICIENCY OF MACHINE LEARNING ALGORITHMS

Nizhny Novgorod State Technical University n.a. R.E. Alekseev

The report is devoted to comparing the effectiveness of machine learning algorithms. The paper considers several machine learning algorithms: polynomial regression; decision tree; random forest. The principles of their work on the basis of the urban transport behavior data are considered.

Keywords: machine learning, decision tree, polynomial regression, random forest.

ИССЛЕДОВАНИЕ ВОЗМОЖНОСТЕЙ СРЕДЫ UNITY ПО ВИЗУАЛИЗАЦИИ РЕЗУЛЬТАТОВ 3D-МОДЕЛИРОВАНИЯ

Нижегородский государственный технический университет им. Р. Е. Алексева

В данной работе проведено исследование возможностей среды Unity по визуализации результатов 3D-моделирования, в том числе, приведены причины необходимости в создании подобного рода приложения. Описаны способы импорта моделей в среду Unity и выбор способа с последующим применением в прототипе приложения.

Ключевые слова: Unity, визуализация, 3D-моделирование, САД, виртуальная реальность.

В наши дни визуализация проникает во все сферы жизнедеятельности людей. 3D-моделирование помогает визуализировать объекты, которые применяются в кинематографе, на телевидении, в компьютерных играх, в медицине и различных областях науки. Это безусловно связано со стремительным развитием новейших технологий, и в результате вместе с ними естественно развивается и улучшается рабочий процесс в различных сферах. До начала новейшей истории пользоваться возможностями 3D-моделирования могли только узкие специалисты: инженеры, трудящиеся в сфере автоматизации и проектирования, занимающиеся математическим моделированием и анализом различных данных. Сегодня доступ к технологиям имеют все команды мира, каждый ученик, любой человек [1]. Командам остается лишь выбрать интересные и полезные функции новых технологий и внедрить их в процесс производства товаров или услуг. 3 отрасли сегодня уже невозможно представить без применения 3D-моделирования [2]:

- индустрия развлечений;
- медицина (хирургия);
- промышленность.

Однако с приходом пандемии COVID-19 пришла необходимость удаленной работы сотрудников [3]. В некоторых компаниях появилась сложность коммуникации при работе над крупными заказами. Раньше, где сотрудники могли подойти к коллеге, чтобы оценить продукт моделирования, сегодня приходится делать много лишних действий [4]. В связи с этим в конструкторских бюро возникла необходимость создания приложения для визуализации результатов 3D-моделирования совместно с коллегами. Для большинства компаний, нуждающихся в подобном решении, оптимальной платформой для создания приложения станет Unity. И этому есть ряд важных причин. Unity – межплатформенная среда разработки, это значит что приложения можно создавать как для ПК, так и для телефонов и интернет-приложений [5]. Unity прост в освоении благодаря модульной системе компонентов и продуманной логике интерфейса. Это значит, что изменять приложение под свои нужды каждая компания сможет без особых затрат. Движок хорошо оптимизирован под работу со средними по размеру 3D-моделями. Есть возможность создания шаблона из любого объекта, поддержка и настройка физики объектов [6]. Наконец, движок поддерживает популярные форматы 3D-моделей и позволяет прикреплять к приложению нужные элементы. Все приведенные выше функции, позволяют компаниям относительно легко настроить приложение под свои нужды. Предпочитаемые САД форматы, с которыми должно работать приложение: X_T, IGES, STEP AP203 и AP214, STL, IPT, PRT, ASM, SAT.

Необходимые функции для проверки:

- возможность загрузки 3D-модели в каждом из доступных форматов в среду unity;
- возможность ручной разборки 3D-модели в каждом из доступных форматов путем выбора отдельных ее частей и их перемещения в указанное пользователем место;
- возможность назначения прозрачности для выделенных частей 3D-модели;
- возможность перемещения 3D-модели в каждом из доступных форматов в указанное пользователем место;
- возможность вращения в трех плоскостях 3D-модели в каждом из доступных форматов;
- возможность отмены всех произведенных над 3D-моделью действий (возврат модели к исходному состоянию);
- возможность создания режима мультиплеер, когда два и более пользователя могут работать с одной моделью;
- возможность добавления моделей в режиме мультиплеер, когда к уже готовой модели любой из участников может добавить другую модель.

Самостоятельно, без сторонних плагинов, Unity способен импортировать модели форматов fbx, dae, 3ds, dxf и obj. При этом документация Unity утверждает, что при установленном на компьютере соответствующих лицензионных приложений Unity способен импортировать форматы приложений Max, Maya, Blender, Cinema4D, Modo и др. [7]. Таким образом только по официальной документации Unity установится понятно, что файлы форматов X_T, IGES, STEP, STL, IPT, PRT, ASM, SAT не читаются Unity самостоятельно. Для их импорта необходимо использовать сторонние плагины, способные читать эти форматы. Анализ источников из интернета привел к 3 плагинам, способным импортировать интересующие нас CAD-форматы. Это плагины:

1. TriLib.
2. CADEXchanger.
3. Pixyz.

Первый вариант способен работать только с форматом STL, не способен импортировать мелкие части модели, только визуальную оболочку. Значит не позволит использовать ручную разборку модели.

Второй вариант не способен импортировать форматы IPT, PRT и ASM.

И только Pixyz поддерживает импорт всех интересующих CAD-форматов.

Для установки плагина в проект Unity необходимо скачать плагин и подтвердить его импорт в проект. При импорте модели формата X_T в проект с установленным Pixyz, Unity сообщает, что файл совместим с плагином и предлагает импортировать модель с его помощью (рис. 1).

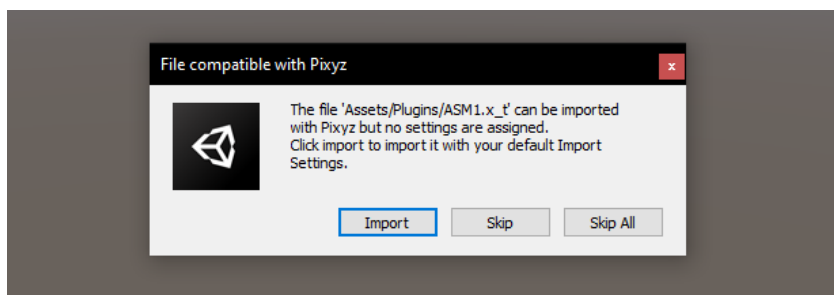


Рис. 1. Предложение Unity импортировать модель формата X_T через Pixyz

На рис. 2 видно, что детали модели, а также её цвета импортируются.

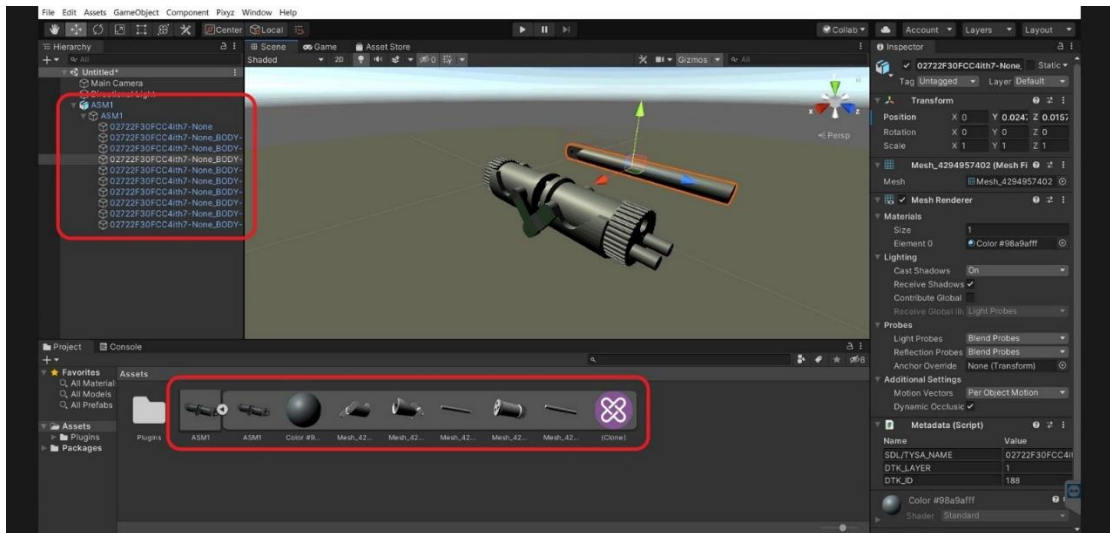


Рис. 2. Результат импорта файла формата X_T с использованием плагина Pixuz

Аналогичные действия были выполнены с остальными САД-форматами. Результаты представлены на рис. 3.

Импорт разных форматов		
ФОРМАТ	ИМПОРТ МЕЛКИХ ЧАСТЕЙ	ИМПОРТ ЦВЕТА
X_T	✓	✓
IGES	✗	✓
STEP 203 и 214	✓	✓
STL	✗	✗
IPT	✓	✓
PRT	✓	✓
ASM	✓	✓
SAT	✓	✓

Рис. 3. Результаты при импорте разных форматов с использованием плагина Pixuz

Возможность ручной разборки модели на сцене определяется самим форматом и тем, как модель экспортировали (если в изначальной модели сохранено разделение модели на мелкие детали, но экспортировали модель как контур, то и Pixuz plugin и Unity распознает в модели только контур). На предыдущем шаге исследования (Исследование возможности загрузки 3D-модели в каждом из доступных форматов в среду Unity) уже было проверено как модели разных форматов с помощью Pixuz импортируются в Unity. Поэтому для простоты восприятия информации, результаты о возможности ручной разборки модели, путем выбора отдельных её частей и перемещения в указанное место, представлены на рис. 4.

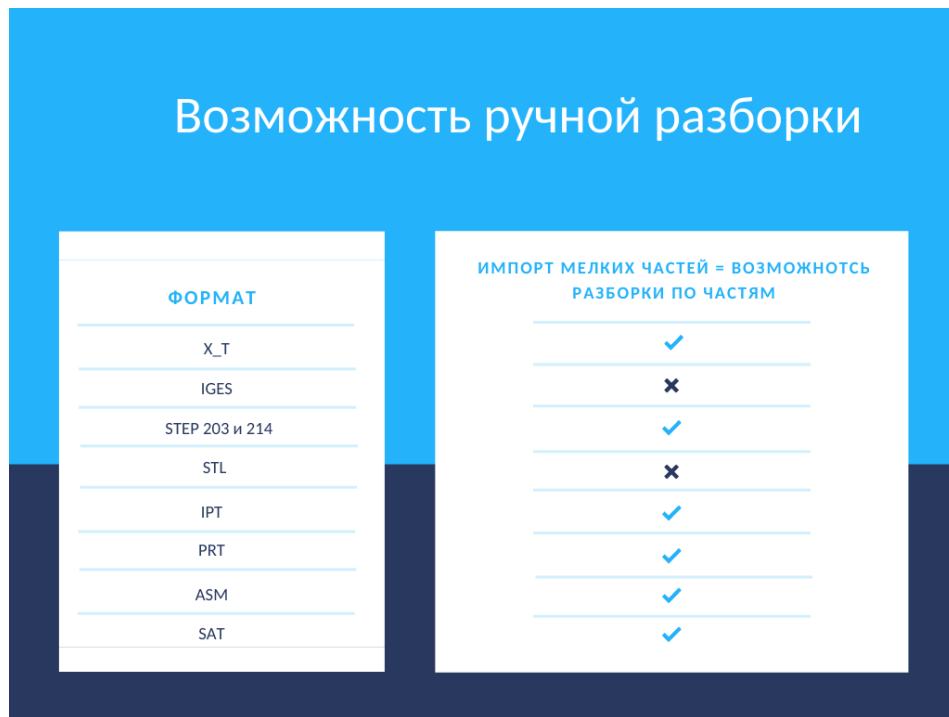


Рис. 4. Возможность ручной разборки в зависимости от формата

В Unity за изменение визуального вида объекта, в том числе и за прозрачность отвечают шейдеры (Shaders). При создании объекта можно создать для него материал, в материале содержится информация в том числе о цвете будущего объекта. Для возможности изменения прозрачности объекта, объект должен иметь материал с параметром Rendering Mode не равный Opaque. Однако Pixuz плагин импортирует модели с Opaque. Это значит, что настройка прозрачности возможна только как полностью непрозрачная модель или часть модели, или как полностью прозрачная.

Перемещение, вращение – одни из базовых действий в Unity. Это значит, что модели, вне зависимости от формата могут быть подвержены этим действиям.

Отмена изменений над моделью достигается благодаря шаблону, который Pixuz автоматически создает для импортированной модели. Таким образом, при желании отменить все действия, которые производились над моделью – Unity просто заменит эту модель на её изначальную форму (шаблон).

За возможность создания мультиплеера в Unity отвечает пакет UNet. Его установка осуществляется посредством установки дополнительного пакета “MultiplayerHLAPI” в проект Unity через окно Window – PackageManager. После настройки по инструкции Unity достигается мультиплеер.

Импорт CAD-моделей в Unity занимается плагин Pixuz, разработчик которого по согласованию с создателями приложения может предоставить возможность импорта моделей на сцену в режиме реального времени. При предоставлении такой возможности второй пользователь может загрузить на сцену другую модель, так как в коде Pixuz нет ограничений на количество моделей на сцене.

Выводы

В данной работе исследованы возможности среды Unity по визуализации результатов 3D-моделирования, приведены причины необходимости в создании подобного рода приложения. Описаны способы импорта моделей в среду Unity и выбор способа с последующим применением в прототипе приложения.

Библиографический список

1. Зубов Н. Развитие 3D-моделирования /Н. Зубов– [Электронный ресурс] / Электронная статья – Россия, 2018 – Режим доступа:<https://seniga.ru/stat/1996-gazvitiye-3d-modjelirovanija.html>, свободный.
2. Романюк А. Трехмерное моделирование в современном мире/А. Романюк. – [Электронный ресурс]/ Электронная статья – Россия, 2019 – Режим доступа: <https://habr.com/ru/post/451266/>, свободный.
3. Конанова М. Прелоадер сайта своими руками /М. Конанова. – [Электронный ресурс] / Электронная статья – Россия, 2020 – Режим доступа: http://www.consultant.ru/document/cons_doc_LAW_348585/, свободный.
4. Пилясова Г. Удаленка – как сохранить непрерывность процессов и коммуникаций/ Г. Пилясова. – [Электронный ресурс] / Электронная статья – Россия, 2020 – Режим доступа: <http://www.garant.ru/article/1363034>, свободный.
6. AppFox. Разработка VR приложений – [Электронный ресурс] / Электронная статья – Россия, 2018 – Режим доступа: <https://appfox.ru/blog/virtualnaya-realnost/razrabotka-vr/>, свободный.
7. Senior UA. Выбор инструментов для создания вашего приложения VR – [Электронный ресурс] / Электронная статья – Россия, 2019 – Режим доступа: <https://senior.ua/articles/vybor-instrumentov-dlya-sozdaniya-vashego-prilozheniya-vr>, свободный.
8. Unitydoc. Как мне импортировать модель из моего 3D редактора? – [Электронный ресурс] / Электронная статья – USA, 2015 – Режим доступа: <https://docs.unity3d.com/ru/530/Manual/HOWTO-importObject.html>, свободный.

A.S. Zimin, U.S. Egorov, N. E. Posobilov

RESEARCH POSSIBILITIES OF THE UNITY ENVIRONMENT ON VISUALIZATION OF THE RESULTS OF 3D MODELING

Nizhny Novgorod state technical university n.a. R.E. Alekseev

In this paper, a study of the capabilities of the Unity environment for visualizing the results of 3D modeling is carried out, including the reasons for the need to create such an application. Described are the ways to import models into the Unity environment and the choice of a method with subsequent application in the application prototype.

Keywords: Unity, visualization, 3D modeling, CAD, virtual reality.

ИССЛЕДОВАНИЕ МЕТОДОВ ВЫБОРА КАРТОГРАФИЧЕСКОЙ ПРОЕКЦИИ В ИНДИКАТОРЕ КРУГОВОГО ОБЗОРА

*Нижегородский государственный технический университет им. Р. Е. Алексеева,
ФНПЦ «ННИИРТ»*

В данном исследовании рассматриваются вопросы применения картографических проекций в индикаторах кругового обзора радиолокационных станций, их значение в отображении радиолокационной информации, также методы и информация, применяемые для определения вида проекций. Результатом проведённого исследования является сформированная теоретическая база, позволяющая учитывать все аспекты формирования и применения проекций, определения их вида, обладающего минимальными искажениями картографируемой области (в зависимости от размера, формы, ориентации и географического положения территории, которую необходимо отобразить). Данные выводы могут быть использованы для моделирования системы выбора оптимальной картографической проекции из большого множества их типов (равновеликие, равноугольные, цилиндрические, азимутальные и т.д.), а также дальнейшей непосредственной ее разработки.

Ключевые слова: радиолокационная станция, индикатор кругового обзора, цифровая карта, система координат, картографическая проекция, искажение, референц-эллипсоид.

В настоящее время радиолокация быстро развивается и получает широкое применение в области обнаружения объектов, их контроля, а также наблюдения за обстановкой в наземном и воздушном пространствах. Инновационные разработки позволяют достигать высокой точности определения координат расположения движущихся аппаратов, следить за их перемещением, выявлять их форму и составляющие части. Процесс конструирования и разработки радиолокационных станций сложен и требует больших вложений сил человека, времени, а также материальных запасов, из чего следует вывод о необходимости автоматизации, выполняющихся в ней процессов. Одним из них является отображение радиолокационной информации на фоне цифровой карты в индикаторе кругового обзора. Для правильного отображения карты, в зависимости от исследуемой ситуации, необходимо выбрать ту картографическую проекцию из их некоторого множества, которая будет вносить минимальную ошибку определения местонахождения(координат) объекта.

В последнее время, существенное развитие получили геоинформационные системы, которые обеспечивают сбор, хранение, обработку и отображение пространственно-координатных данных. Данные системы позволяют работать с тематическими и топографическими картами, различными системами координат, картографическими проекциями. В связи с этим, они предоставляют интерфейс, позволяющий выбрать необходимую проекцию, однако при этом оператор радиолокационной станции изначально должен знать, какой тип проекции ему необходим, что в условиях использования станции является практически невозможным (мобильность станции, недостаточная квалификация оператора, необходимость быстрого принятия решения). Следствием указанных факторов является необходимость в разработке методики, с дальнейшей её автоматизацией, подбирающей тип проекции цифровой карты для отображения на индикаторе кругового обзора. Данная работа направлена на рассмотрение методов применения картографических проекций, а также различных факторов, которые оказывают влияние на принимаемый вид проекции.

Индикаторные устройства, входящие в состав радиолокационной станции, обеспечивают возможность, посредством графических интерфейсов, отображать радиолокационную информацию, получаемую в результате приема и обработки отражённых от объектов сигналов [1, 2]. Радиолокационная информация на экране индикатора отображается в полярных координатах азимут-дальность на фоне цифровой карты местности. Карта содержит цифровую информацию о земной поверхности, которая необходима для работы станции. Подобная информация обеспечивает решение следующих основных задач:

- изучение и оценка свойств и параметров местности;
- определение координат стационарных и движущихся объектов в реальном масштабе времени;
- решение расчётных задач и наглядное отображение полученных результатов.

Конечной целью всех вышеперечисленных задач является максимально точное определение положения геодезических объектов на поверхности Земли, изображаемой в форме принятого референц-эллипсоида. Положение этих объектов может быть определено в различных системах координат. Однако все процедуры для выполнения заданий ведутся в геодезической и плоской прямоугольной системах, которая в свою очередь является более удобной для практического использования. Но их применение вместо геодезических координат требует перехода от поверхности эллипсоида на плоскость, что выполняется за счет использования картографических проекций, которые порождают различного рода искажения [3].

Под картографической проекцией понимается некоторый определённый математический закон отображения, устанавливающий однозначное и непрерывное соответствие между точками поверхности эллипсоида (сферы) и плоскости. Это соответствие может быть задано уравнениями вида (1) [8]:

$$X = f_1(\varphi, \lambda); Y = f_2(\varphi, \lambda) \quad (1)$$

где X, Y – координаты точки на плоскости, φ, λ – координаты точки на поверхности эллипсоида (широта и долгота), f_1, f_2 – некоторая функциональная зависимость.

Картографическая проекция выбирается из множества видов проекций, она должна обеспечивать минимум искажений. Искажения проекций бывают нескольких видов:

- искажение длин: подразумевает непостоянство масштаба плоского изображения. Это означает, что в различных точках поверхности или в одной точке, но различных направлениях, будет наблюдаться изменение масштаба;
- искажение площадей: характеризуется отклонением площади эллипса искажений от исходной площади на эллипсоиде.
- искажение углов трактуется как разность углов между направлениями на карте и соответствующими направлениями на поверхности эллипсоида.
- искажения формы – это графическое изображение вытянутости эллипсоида.

Факторы, влияющие на выбор картографической проекции различны, и могут быть условно разделены на несколько групп.

К первой группе относятся факторы, характеризующие объект картографирования: географическое положение изображаемой территории, её размеры, форму, значимость и степень показа смежных территорий.

Вторая группа включает факторы, характеризующие создаваемую карту, способы и условия её использования: назначение и специализацию, масштаб и содержание карты, задачи, которые будут решаться по ней и требования к точности их решения, способы анализа картографической информации.

К третьей группе относятся факторы, которые характеризуют получаемую картографическую проекцию: характер искажений, условия обеспечения минимума искажений и допустимые максимальные искажения длин, углов и площадей, характер их

распределения, кривизна изображения геодезической линии, условия изображения других линий положения, степень передачи форм территории, кривизна изображения линий картографической сетки, характер изображения полюсов и т.п. [4].

Выбор картографических проекций осуществляется в два этапа:

1. сначала определяют совокупность проекций, которые наиболее подходят для отображения составляемых карт;
2. для сформированного множества проекций оценивают искажения, вносимые картографической проекцией, и выбирают минимальное значение искажений по всем показателям - выполняют окончательный выбор.

Факторы выбора проекции, относящиеся к первой группе, считаются твёрдо заданными и предполагают выбор таких проекций, в которых их центральные точки и центральные линии, вблизи которых масштабы мало изменяются, находятся в центре картографируемой территории, а центральные линии располагаются по направлению наибольшего распространения этих территорий.

Вторая группа факторов является основной при выборе картографических проекций, на основе которой определяют относительную значимость факторов третьей группы, т.е. какие из требований в данном конкретном случае наиболее существенны, а какие из них можно не учитывать. К числу строго учитываемых условий относят: желаемый характер искажений проекции, максимально допустимые искажения, изображение полюсов, симметричность или асимметричность картографической сетки. В соответствии со значениями данных параметров, выбираются те проекции, в которых они выполняются. Таким образом, факторы, приобретающие в данном конкретном случае безусловную значимость, в дополнение к факторам первой группы позволяют в основном решить первую часть задачи.

Выполнение тех или других ограничительных требований к математическим элементам карт очень часто влечёт за собой увеличение искажений в целом [5-6]. Однако, характерными требованиями, которые должны соблюдаться, являются следующие:

- полное отсутствие искажения углов или площадей на всей карте или в некоторых её местах;
- изображение меридианов и/или параллелей, в виде отрезков прямых, дуг окружностей или эллипсов;
- соблюдение симметрии сетки большей частью относительно экватора и среднего меридиана.

Исходя из выше перечисленных факторов, которые приобретают некоторую значимость в зависимости от конкретной ситуации, выбирается определённая проекция, которая является частью одной из нижеперечисленных групп и подразумевает некоторые специфические свойства [7-9]:

1. Группы по характеру искажений:
 - равноугольные проекции – проекции без искажений углов, в которых масштаб зависит только от положения точки и не зависит от направления;
 - равновеликие проекции – подразумевают отсутствие искажения площадей, но подвержены сильному искажению углов и форм;
 - **равнопромежуточные** проекции – каждая точка карты обладает свойством сохранения длины по одному из главных направлений.
2. Группы по виду параллелей и меридианов:
 - круговые проекции: меридианы и параллели изображаются окружностями, экватор и средний меридиан – прямые линии. Применяются для изображения всей поверхности Земли.
 - цилиндрические проекции – параллели и меридианы изображаются двумя семействами параллельных прямых линий, перпендикулярных друг другу. Цилиндрические проекции применяются при составлении карт мелких и крупных масштабов (рис. 1).

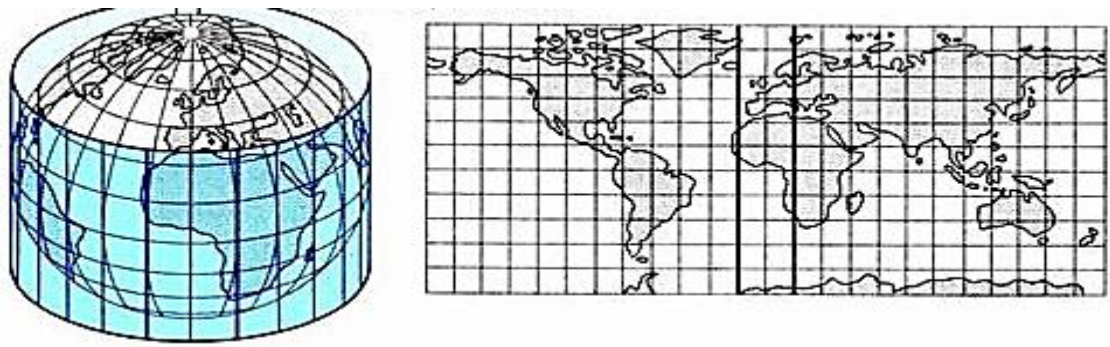


Рис. 1. Цилиндрическая проекция

- конические проекции – координатные линии Земли проектируют по какому-либо из законов на внутреннюю поверхность описанного или секущего конуса, а затем, разрезав конус по образующей, разворачивают его на плоскость (рис. 2).

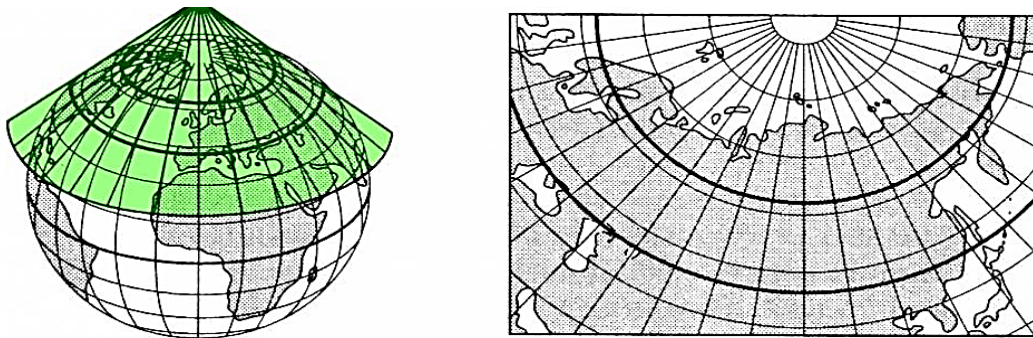


Рис. 2. Коническая проекция

- азимутальные проекции – (рис.3) параллели изображаются концентрическими окружностями с центром в полюсе, а меридианы – пучком прямых, исходящих из центральной точки проекции, углы между меридианами проекции равны соответствующим разностям долгот.

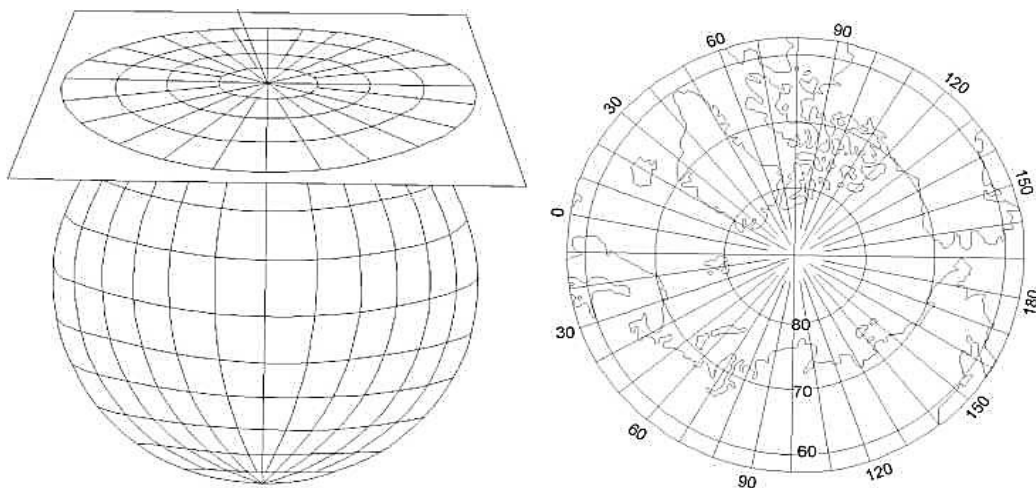


Рис. 3. Азимутальная проекция

Подводя итог проведённой работы, можно сделать вывод, что выбор картографической проекции - сложный процесс, требующий учёта большого количества различных факторов. В связи с этим, данный процесс необходимо начинать с анализа ранее созданных аналогичных карт и их математической основы, а затем проекций, рекомендуемых в атласах для выбора картографических проекций. Кроме того, в случаях постоянного использования, рекомендуется его автоматизация, путём составления алгоритмической и математической основы с помощью прикладных и программных средств.

Библиографический список

1. Панасюк Ю.Н. Обработка радиолокационной информации в радиотехнических системах: учебное пособие / Ю.Н. Панасюк, А.П. Пудовкин. – Тамбов: Изд-во ФГБОУВПО «ТГТУ», 2016. – 84 с.
2. Фомин А.Н. Теоретические и физические основы радиолокации и специального мониторинга: учебник / А.Н. Фомин, В.Н. Тяпкин, Д.Д. Дмитриев [и др.]; под общ. ред. И. Н. Ищука. – Красноярск: Сиб. федер. ун-т, 2016. – 292 с.
3. Вильданов И.Р., Мухаметдинов А.Р. Картографические проекции // Заметки учёного. – 2017. – №3(19). – 31-33.
4. Ахметзянова А.И. Радиолокационные методы обнаружения объектов // Международная научно-практическая конференция «САПР и моделирование в электронике». – 2018.
5. Загребин Г.И. Разработка атласа картографических проекций на крупные регионы Российской Федерации // Известия высших учебных заведений. Геодезия и аэрофотосъёмка. – 2012. – №2. – 52-55.
6. Топчилов М.А. Аналитическое исследование свойств картографических проекций: метод. указания по выполнению контрольной работы для студентов заочного факультета / М.А. Топчилов, Л.А. Ромашова, О.Н. Николаева. – Новосибирск, СГГА. – 2006. – 25 с.
7. Бажукова Н.В. Картография: учебное пособие / сост. Н.В. Бажукова. – Пермский государственный национальные университет. – Пермь, 2020, 310с.
8. Витковский В. Картография (теория картографических проекции) / В. Витковский. – М.: Нобель Пресс, 2013. – 574 с.
9. Утешева Т.Ш. Математическое проецирование в ГИС: учебно-методическое пособие / Т.Ш. Утешева. – Нижний Новгород: "Нижегородский госуниверситет", 2016. – 60 с.

К.О. Staroverova, D.V. Kiselev

INVESTIGATION OF METHODS FOR SELECTING A MAP PROJECTION IN THE CIRCULAR VIEW INDICATOR

*Nizhny Novgorod state technical university n.a. R.E. Alekseev
FRPC NNIIRT*

This study deals with the application of cartographic projections in the indicators of the radar stations circular view, their significance in the display of radar information, as well as the methods and information used to determine their type. The result of the conducted research is the formed theoretical base that allows to take into account all aspects of the formation and application of projections and correctly make decisions in various situations (depending on the location of the radar station, the size of the territory to be displayed and their purpose). These conclusions can be used to model the system for selecting a map projection, as well as to further develop it directly.

Keywords: radar station, circular view indicator, digital map, coordinate system, cartographic projection, distortions, reference ellipsoid.

ПОДХОД К КОНЦЕПТУАЛЬНОМУ ПРОЕКТИРОВАНИЮ ИННОВАЦИОННЫХ ИНФОРМАЦИОННЫХ СИСТЕМ

Нижегородский государственный технический университет им. Р.Е. Алексеева

Работа посвящена разработке подхода к концептуальному проектированию инновационных информационных систем, которые разрабатываются, как правило, в условиях неполноты и существенной неопределенности относительно потребностей пользователей и востребованности результатов разработки. Развиваемый подход позволит повысить полезность и качество разрабатываемых информационных систем и сервисов на основе синтеза модели полезности и процедур валидации.

Ключевые слова: концептуальное проектирование, информационная система, поддержка принятия решений, качество программного обеспечения.

Сложившаяся практика показывает, что соответствие разрабатываемых инновационных информационных систем (и сервисов) пользовательским, функциональным и нефункциональным требованиям спецификаций и технического задания не гарантирует, что программный продукт будет востребован у конечных пользователей, т.е. фактическая полезность информационных систем может оказаться ниже ожидаемой. Такое положение дел приводит, как правило, к низким показателям продаж и, как следствие, сворачиванию проекта.

Прогнозирование полезности инновационных информационных систем (и сервисов) является сложной задачей ввиду того, что на рынке слабо представлены конкурирующие программные продукты, что затрудняет сравнение. Зачастую, у разрабатываемого программного продукта вообще нет аналога, т.к. в основе лежит инновационная технология, опыт применения которой еще не сложился.

Известным примером является продвижение высокотехнологичной продукции компании Apple на заре развития сенсорных экранов для мобильных телефонов. Предлагаемая технология была инновационной для того времени, а мнения экспертов в предметной области были полярно противоположными. Часть экспертов предсказывала провал и банкротство компании, другая часть говорила о том, что продукт будет полезен для пользователей.

В рамках работы под полезностью понимается основная характеристика набора функций (конфигурации) инновационного продукта, необходимая для удовлетворения определенных потребностей пользователей.

Выявление и формализация потребностей пользователей является частью процесса управления требованиями, одного из начальных этапов разработки инновационного ИТ-проекта.

Инновационный ИТ-проект – это набор взаимосвязанных процессов, состоящих из скоординированных и управляемых задач с начальной и конечной датами, предпринятых для создания инновационного (уникального) ИТ-продукта или ИТ-услуги.

К методам выявления потребностей пользователей относятся: интервьюирование и анкетирование; проведение семинаров; мозговой штурм; работа в фокус-группами и др. Как правило, выявлением потребностей пользователей занимается профильный специалист по требованиям.

Одним из вариантов формализации пользовательских потребностей являются варианты использования (или пользовательские сценарии), на основе которых затем формулируются пользовательские требования.

Пользовательские требования описывают функциональные и нефункциональные требования к программному продукту так, чтобы они были понятны даже пользователю, не имеющему специальных технических знаний. Эти требования определяют только внешнее поведение системы, избегая определения структурных характеристик системы.

Функциональные же требования определяют функциональность программного обеспечения (ПО), которую разработчики должны обеспечить, чтобы пользователи смогли выполнить свои задачи в рамках пользовательских требований.

В итоге пользовательские и функциональные требования отражаются в техническом задании (ТЗ) на разработку программного продукта наряду с техническими характеристиками, показателями качества и технико-экономическими требованиями, состава требуемой документации, а также некоторые специальные требования.

Особенность разработки инновационных проектов заключается в том, что конечный программный продукт, отвечающий всем установленным требованиям ТЗ, может не соответствовать целям и ожиданиям пользователей.

С одной стороны, конечный продукт, выпущенный на рынок, проходит верификацию, т.е. внутренний процесс управления качеством, обеспечивающий подтверждение того факта, что продукт полностью соответствует всем установленным ТЗ требованиям. С другой стороны, валидация – процесс приведения доказательств того, что потребности конкретного пользователя продукта или услуги удовлетворены; выполняется, как правило, в условиях либо опытной эксплуатации, либо рыночной конкуренции. При этом, если программный продукт не обладает «ценностью» или полезностью для пользователей, то он не будет коммерчески успешным.

Под качеством понимается степень удовлетворения пользователя конечным программным продуктом, однако, при разработке инновационных продуктов, это заранее неизвестно, поэтому принятие решений относительно конфигурации функций программного продукта происходит в условиях неопределенности.

Все это предопределяет необходимость выработки подхода, который позволит повысить эффективность управления отдельными процессами инновационных проектов в ИТ-сфере для создания продукта, востребованного среди пользователей.

Цель работы состоит в том, чтобы повысить полезность проектируемых инновационных информационных систем (и сервисов) на основе совершенствования процедур валидации и синтеза иерархической модели полезности (для пользователей).

После завершения этапа выявления требований, в ТЗ на разработку ПО определен состав функций, которые должны быть реализованы в итоговом программном продукте.

Предполагается, что не все определенные в ТЗ функции окажутся полезными для конечных пользователей, поэтому, прежде чем приступать к проектированию, необходимо найти такое концептуальное решение (набор функций), которое при максимальной полезности конечного продукта обеспечивает минимальные проектные и технические риски, и удовлетворяет ограничением по трудоемкости.

Поскольку условием неопределенности считается ситуация, когда последствия принимаемых решений лицом, принимающим решения (ЛПР), неизвестны, и можно лишь приблизительно их оценить, то для принятия решения можно использовать разные критерии, позволяющие найти наилучшее решение, в общем случае максимизирующее возможную прибыль и минимизирующее возможный убыток.

Одним из возможных критериев при этом является критерий Сэвиджа, который также известен как «Критерий минимальных сожалений», также можно рассматривать его с обратной стороны (с точки зрения пользователей), то есть определять решение максимизирующее ожидаемую полезность по всем возможным конфигурациям ИС.

В предлагаемом же подходе предлагается определять ценность разрабатываемого программного обеспечения, учитывая при этом:

- мнение экспертов, которые могут оценивать вероятность реализации той или иной функции и их потенциальную полезность;
- мнение пользователей (например, фокус-группы), которые могут оценить полезность тех функций, которые ранее были реализованы в других «похожих» продуктах (если они существуют);
- трудоемкость реализации каждой функции, что является ограничением, которое позволит сократить возможное количество функций в разработке.

Таким образом, оценивая показатели (полезность U , вероятность P и трудоемкость T реализации заявляемых функций), можно сформировать множество альтернатив (конфигураций ПО).

Чтобы сравнивать альтернативы необходимо определить целевую функцию, при этом выбор наилучшего варианта может осуществляться путем сведения многокритериальной задачи выбора к однокритериальной с использованием дополнительной информации. Перейти к однокритериальной задаче позволяют методы свертки векторного критерия и методы ранжирования критериев.

В таблице 1 представлено множество вариантов решения задачи, где 1 – показатель является целевой функцией, 0 – показатель является ограничением.

Таблица 1

Варианты решения задачи

Показатели	Вар. 1	Вар. 2	Вар. 3	Вар. 4	Вар. 5	Вар. 6	Вар. 7	Вар. 8
U	0	1	0	0	1	1	0	1
T	0	0	1	0	1	0	1	1
P	0	0	0	1	0	1	1	1

После выбора одного из восьми возможных вариантов, необходимо, чтобы эксперты(разработчики) оценили вероятность реализации каждой функций ТЗ, а пользователи (фокус-группа) их потенциальную полезность.

Суммарную предельную трудоемкость реализации можно считать всегда известной величиной, которую нельзя превышать. Минимальная вероятность успешности проекта тоже считается заданной исходными данными к проекту.

Рассмотрим пример создания инновационного программного продукта, в архитектуре которого выделено 5 модулей, реализующих заданную ТЗ функциональность с вероятностью успешности проекта не ниже 0,7. Успешным считается проект, который проходит верификацию. Пусть трудоемкость реализации каждой функции известна и представляет собой количество нормо-часов разработки, а общая трудоемкость на проект 400 н/час. Эксперты оценивают вероятность успешной реализации каждой функции от 0 до 1, а пользователи полезность (важность) от 1 до 5.

Исходные данные представлены в таблице 2.

Таблица 2

Исходные данные

Модули	U	T	P
Функция 1	5	210	0,8
Функция 2	3	130	0,9
Функция 3	4	170	0,9
Функция 4	3	50	0,9
Функция 5	4	90	0,9

Для 5 функциональных модулей возможны 32 конфигурации ПО (табл. 3).

Возможные конфигурации ПО

№	Функция 1	Функция 2	Функция 3	Функция 4	Функция 5	U	T	P
0	0	0	0	0	0	0	0	0
1	0	0	0	0	1	4	90	0,90
2	0	0	0	1	0	3	50	0,90
3	0	0	0	1	1	7	140	0,81
4	0	0	1	0	0	4	170	0,90
5	0	0	1	0	1	8	260	0,81
6	0	0	1	1	0	7	220	0,81
7	0	0	1	1	1	11	310	0,73
8	0	1	0	0	0	3	130	0,90
9	0	1	0	0	1	7	220	0,81
10	0	1	0	1	0	6	180	0,81
11	0	1	0	1	1	10	270	0,73
12	0	1	1	0	0	7	300	0,81
13	0	1	1	0	1	11	390	0,73
14	0	1	1	1	0	10	350	0,73
15	0	1	1	1	1	14	440	0,66
16	1	0	0	0	0	5	210	0,8
17	1	0	0	0	1	9	300	0,72
18	1	0	0	1	0	8	260	0,72
19	1	0	0	1	1	12	350	0,65
20	1	0	1	0	0	9	380	0,72
21	1	0	1	0	1	13	470	0,65
22	1	0	1	1	0	12	430	0,65
23	1	0	1	1	1	16	520	0,58
24	1	1	0	0	0	8	340	0,72
25	1	1	0	0	1	12	430	0,65
26	1	1	0	1	0	11	390	0,65
27	1	1	0	1	1	15	480	0,58
28	1	1	1	0	0	12	510	0,65
29	1	1	1	0	1	16	600	0,58
30	1	1	1	1	0	15	560	0,58
31	1	1	1	1	1	19	650	0,53

Выбираются наиболее полезные конфигурации, которые не превышают трудоемкость (400 н/час) и имеют вероятность не ниже заданной (0,7), результаты расчета также приведены в табл. 3. Из всех оставшихся вариантов выбирается конфигурация с максимальной полезностью, минимальной трудоемкостью и максимальной вероятностью успешной реализации, а именно, в примере выбрана конфигурация № 7.

Альтернативы можно сравнивать между указанным по следующим показателям:

- экономия – полезность (рис. 1);
- вероятность реализации – полезность (рис. 2);
- экономия – вероятность (рис. 3).

Под экономией понимается возможность сэкономить бюджет за счет снижения трудоемкости реализации проекта.

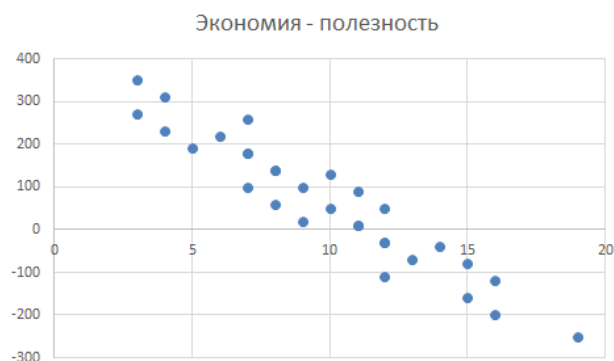


Рис. 1. Сравнение альтернатив по основаниям «Экономия – Полезность»

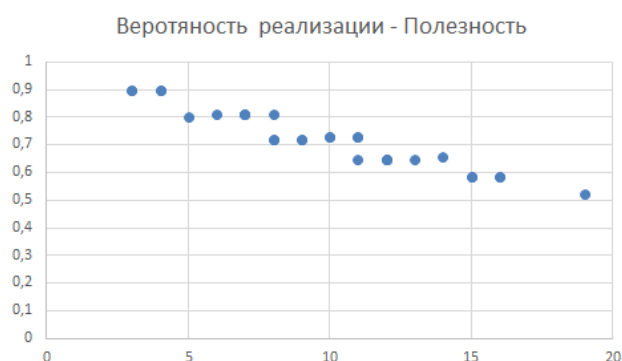


Рис. 2. Сравнение альтернатив по основаниям «Вероятность реализации – Полезность»



Рис. 3. Сравнение альтернатив по основаниям «Экономия – Вероятность реализации»

В результате ЛПР в процессе концептуального проектирования сможет более взвешено принимать решения относительно конфигурации функций программного продукта в условиях неопределенности, что позволит повысить полезность разрабатываемых ИС и сократить трудоемкость реализации ПО.

Библиографический список

1. Баранов В.Г., Седаков Д.В., Милов В.Р. Способ управления конфигурацией информационных систем // Информационно-измерительные и управляющие системы. 2017. № 8 (15). С. 26-33.
2. Скрипник И.В., Егоров Ю.С. Выбор системно-технических решений при построении радиосвязного оборудования с применением аппарата нечеткой логики // Проектирование и технология электронных средств. 2019. № 3. С. 21-27.

Y.S. Yegorov

**APPROACH TO THE CONCEPTUAL DESIGN OF INNOVATIVE
INFORMATION SYSTEMS**

Nizhny Novgorod state technical university n.a. R.E. Alekseev

The paper considers an approach to the conceptual design of innovative information systems, which are developed in conditions of incompleteness and significant uncertainty regarding the needs of users and the demand for development results. The approach will improve the usefulness and quality of the developed information systems based on the synthesis of the utility model and validation procedures.

Keywords: conceptual design, information system, decision support, software quality.

ПОДХОД К РАЗРАБОТКЕ ИНТЕЛЛЕКТУАЛЬНОЙ СОВЕТУЮЩЕЙ СИСТЕМЫ ПОДДЕРЖКИ СТИЛИСТИЧЕСКИХ РЕШЕНИЙ С ПРИМЕНЕНИЕМ НЕЧЕТКОЙ ЛОГИКИ

Нижегородский государственный технический университет им. Р. Е. Алексеева

В работе рассматривается подход к описанию и выбору стилистических решений для подбора наиболее актуальных вариантов одежды и их комбинаций. Подход позволяет формализовать требования с применением аппарата нечеткой логики и определить частные целевые функции для поддержки принятия стилистических решений.

Ключевые слова: поддержка принятия решений, информационная система, дизайн, стилистические решения, нечеткая логика.

Системы поддержки принятия решений (СППР) – это интерактивные системы, используемые в управлении для помощи в принятии решений в условиях частично и слабо структурированных задач. Они имеют развитый аналитический аппарат по сравнению с управленческими информационными системами, а также включают некоторую совокупность математических моделей. С целью реагирования на внешние изменяющиеся условия СППР в качестве входных данных используют информацию из офисных, профессиональных и транзакционных систем и систем управления, а также должны получать информацию из внешних источников.

В рамках разработки интеллектуальной советуемой системы поддержки стилистических решений предполагается, что на основе запроса пользователя определяется его стилистический образ, информация о котором в дальнейшем будет использоваться в процессе интерактивной визуализации. Под интерактивной визуализацией понимается организация обратной связи таким образом, что предлагаемые системой советы будут актуальны для пользователя в рамках жизненной ситуации. Например, такой ситуацией можно считать собеседование соискателя на работу, когда внешний вид во многом влияет на общее впечатление и итоговый результат.

По данным компании HeadHunter 82% соискателей детально продумывают свой внешний вид перед собеседованием. Известно, что первое впечатление о человеке формируется в первые несколько секунд, а когда у рекрутера только резюме, то хорошее впечатление можно создать за счет внешнего вида. Описывая идеального соискателя, каждый третий менеджер по персоналу называл официальный стиль: костюм неяркого цвета (пиджак и брюки классического кроя) плюс строгая сорочка. Для женщин допустима юбка длиной до середины колена и сорочка с рукавом «три четверти». Обувь должна быть неброской, для женщин – обязательно закрытой, на небольшом каблуке.

Таким образом, способ анализа стилистических решений на основе нечеткого моделирования требований может составить основу для интеллектуальной советуемой системы.

На этапе создания образа сначала создается абстрактная модель, затем формируется одно или несколько потенциальных стилистических решений. Для сопоставления нескольких вариантов стилистических решений необходимо определить критерии сравнительной оценки этих решений. При этом важно, чтобы все исходные требования, содержащиеся в задании пользователя на подбор образа, были учтены, а предлагаемые решения полностью адекватны.

Обозначим совокупность требований, к характеристикам n -го компонента стилистического образа как R_{nl} , $l = \overline{1, L_n}$. При выборе стилистических решений требования R_{nl} , связанные с числовыми характеристиками U_{nl} , формализуются

заданием интервалов значений в виде односторонних неравенств $x_{nl} \geq x_{Lnl}$ или $x_{nl} \leq x_{Hnl}$ либо двусторонних неравенств $x_{Lnl} \leq x_{nl} \leq x_{Hnl}$. Допустимые значения дискретных характеристик могут быть заданы равенством либо перечислением с использованием логических связок \wedge – конъюнкция и \vee – дизъюнкция. Так, если образ состоит из рубашки, костюма и галстука, а при этом необходимо поддерживать одну однотонную вещь и две с рисунком, то соответствующее требование выполняется, если одновременно выполняются все элементарные «дочерние» требования. Другое требование может заключаться в поддержке двух однотонных вещей и одной с рисунком [1]. Для учета выполнения требований введем индикаторную (логическую) переменную r_{nl} такую, что $r_{nl} = 1$, если требование R_{nl} выполняется и $r_{nl} = 0$ в противном случае.

Представим результат проверки выполнения l -го требования, k -го варианта реализации, n -го компонента образа в виде

$$r_{nlk} = F_{R_{nl}}(x_{nlk}), \quad (1)$$

где $F_{R_{nl}}(\bullet)$ – формальная спецификация требований, x_{nlk} – значение l -й характеристики, k -го варианта реализации, n -го компонента образа.

Для выбора варианта стилистического решения предлагается следующий алгоритм. Сначала на основе анализа требований пользователя для образа формируются ограничения, которые определяют множество допустимых решений V_{An} .

Затем на основе проверки выполнения требований для отдельных вариантов образа (ВР) v_{nk} , $k = \overline{1, K}$ последовательно проверяется выполнение требований. В результате для каждого ВР v_{nk} проверяется принадлежность множеству допустимых решений V_{An} . При этом

$$v_{nk} \in V_{An}, \text{ если } r_{nk} = 1, \text{ где } r_{nk} = \bigwedge_{l=1}^{L_n} r_{nlk}. \quad (2)$$

Если множество допустимых решений V_{An} содержит один единственный элемент, то этот вариант является решением \hat{v}_{nk} задачи. Однако этот случай достаточно редко возникает на практике. Чаще множество допустимых решений V_{An} содержит несколько элементов либо оказывается пустым.

Если множество допустимых решений содержит несколько элементов, то выполняется процедура выбора одного из нескольких допустимых вариантов реализации. Поскольку индикаторная переменная выполнения требований $r_{nk} = 1$ для всех допустимых ВР $v_{nk} \in V_{An}$, то для их сравнения ВР предлагается применять аппарат нечеткой логики.

Так для формализации требований вместо указания пороговых значений характеристик с определением двух градаций выполнения требований (выполняется и не выполняется), предлагается задавать требования с помощью функций принадлежности (ФП) нечеткому множеству $r = \mu_r(x)$. Степень выполнения модифицированного требования R определяется, как значение соответствующей ФП $r \in [0; 1]$.

В качестве основных вариантов формализации требований, связанных с количественными характеристиками компонент образа, предлагается использовать в развитие [2]: линейную s -образную ФП для правосторонних ограничений, линейную z -образную ФП для левосторонних и трапецевидную ФП для двухсторонних ограничений.

Поэтому ФП для формализации требований [3] можно рассматривать как переход от жестких ограничений с одним порогом к двухпороговым мягким ограничениям.

Для учета особенностей выполнения требований предложены функции принадлежности (рис. 1) к нечеткому множеству специального вида:

$$\mu_s(x) = \mu_s(x; x_B, x_T, c) = \begin{cases} 0 & ; x < x_B \\ c & ; x = x_B \\ \frac{(1-c)x + cx_T - x_B}{x_T - x_B} & ; x_B < x \leq x_T \\ 1 & ; x > x_T \end{cases} \quad (3)$$

– для представления односторонних ограничений снизу;

$$\mu_z(x) = \mu_z(x; x_B, x_T, c) = \begin{cases} 1 & ; x < x_B \\ \frac{(c-1)(x-x_B)}{x_T - x_B} & ; x_B < x \leq x_T \\ 0 & ; x > x_T \\ c & ; x = x_T \end{cases} \quad (4)$$

– для представления односторонних ограничений сверху.

Параметрами ФП (3) и (4) являются граничные требуемые значения, определяемые исходными ограничениями-неравенствами, дополнительное значение, начиная с которого ФП принимает единичное значение и оценка степени выполнения требования c при достижении граничного требуемого значения. «Ступенька» в ФП (рис. 1) обеспечивает переход через точку, соответствующую граничному требуемому значению.

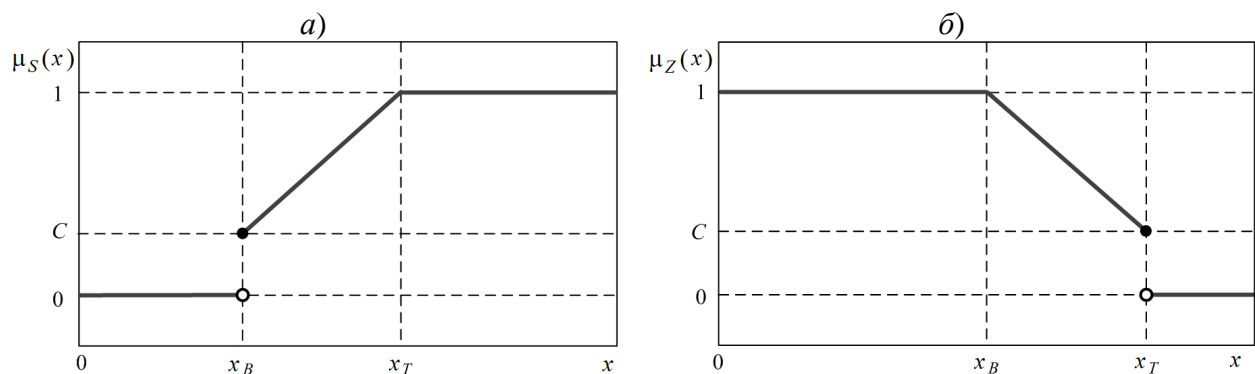


Рис. 1. График функции принадлежности: s-вида (а) и z-вида (б)

Для примера рассмотрим характеристики (табл. 1) вариантов стилистического образа.

Поскольку все требования выполняются только для ВР-1, то согласно описанному выше алгоритму, этот первый вариант реализации \hat{v}_{n1} может быть принят в качестве искомого. Если учесть, что первое и четвертое требования опциональные, то третий вариант реализации v_{n3} также будет принадлежать множеству допустимых решений V_{An} . В этом случае возникает необходимость в дальнейшем сравнении двух альтернатив $v_{n1} \in V_{An}$ и $v_{n3} \in V_{An}$.

Таблица 1

Характеристики ВР

Требования	Варианты реализации
------------	---------------------

№	Характеристики	Целевые хар-ки	ВР-1	ВР-2	ВР-3
1	Куртка	Черный (опция); XL (об.); Покрой (об.)	Черный; XL; Прямой	Черный; XL; Прямой	Белый; XL; Прямой
2	Майка (% хлопка)	Не менее 20	30	30	25
3	Пиджак (кол-во карманов, шт.)	Не более 3	2	2	3
4	Жилетка	Опция	Есть	Есть	Нет
5	Классические брюки	Черный (об.)	Черный	Черный	Черный
6	Рубашка	Белый (об.)	Белый	Розовый	Белый
7	Носки (высота от щиколотки, см)	Не менее 1,5	2	2	3
8	Аксессуары, шт.	Не более 8	6	8	5

Таблица 2

Выполнение требований для ВР

№	Требования	Опция	Варианты реализации		
			ВР-1	ВР-2	ВР-3
1	R1: $y_1 = a_{1,1} \wedge a_{1,2} \wedge a_{1,3}$	*	1	1	0
2	R2: $y_2 \geq 20$		1	1	1
3	R3: $y_3 \leq 3$		1	1	1
4	R4: $y_4 = a_4$	*	1	1	0
5	R5: $y_5 = a_{5,1} \vee a_{5,2}$		1	1	1
6	R6: $y_6 = a_6$		1	0	1
7	R7: $y_7 \geq 1,5$		1	1	1
8	R8: $y_8 \leq 8$		1	1	1

Для этого можно применить двухпороговые мягкие ограничения и найти параметры ФП (3) и (4). Для ФП s -вида (рис.1, а) параметры определяются выражением

$$x_{Вnl} = x_{Lnl}, \quad x_{Тnl} = \max_{k=1,K} x_{nk}. \quad (5)$$

Для ФП z -вида (рис.2, б) параметры ФП определяются выражением

$$x_{Вnl} = \min_{k=1,K} x_{nk}, \quad x_{Тnl} = x_{Hnl}. \quad (6)$$

Оценку степени выполнения требования при достижении граничного требуемого значения примем равной $c = 0,5$. Обоснование выбора этого значения может составить предмет дополнительного исследования.

Для опциональных требований задаются весовые коэффициенты $0 < w_{nl} < 1$, характеризующий значимость отдельных требований. При нулевом значении весового коэффициента, соответствующее требование исключается из рассмотрения.

Результат анализа выполнения требований (табл. 1) для двух отобранных ВР-1 и ВР-3 на основе предложенного подхода, с применением аппарата нечеткой логики представлен в табл. 3.

Таблица 3

Оценки степени выполнения модифицированных требований для ВР

№	Функции принадлежности	Весовой коэффициент	Варианты реализации			
			Без учета весов		С учетом весов	
			ВР-1	ВР-3	ВР-1	ВР-3
1	$\mu_{R1}(y)$	0,4	1	0	1	0,6
2	$\mu_{R2}(y) = \mu_S(y; 20; 30; 0,5)$	1	1	0,75	1	0,70
3	$\mu_{R3}(y) = \mu_Z(y; 2; 3; 0,5)$	1	1	0,5	1	0,5
4	$\mu_{R4}(y)$	0,6	1	0	1	0,4
5	$\mu_{R5}(y)$	1	1	1	1	1
6	$\mu_{R6}(y)$	1	1	1	1	1
7	$\mu_{R7}(y) = \mu_S(y; 1,5; 3; 0,5)$	1	0,833	1	0,833	1
8	$\mu_{R8}(y) = \mu_Z(y; 5; 8; c)$	1	0,75	1	0,75	1

При определении параметров ФП коэффициенты соответствия требованиям характеризуют близость альтернатив (ВР) к множеству допустимых решений.

При решении такой задачи, в отличие от процедуры выбора на множестве допустимых решений, возникает необходимость в дополнительной оценке для различных ВР вероятности достижения целей пользователя.

Определение вероятности достижения целей пользователя обычно выполняется с привлечением известных процедур обработки экспертных мнений и выходит за рамки данной работы.

Предложенный способ анализа стилистических решений на основе нечеткого моделирования требований может составить основу для советующей системы поддержки стилистических решений. Формализация требований с применением аппарата нечеткой логики позволяет сравнивать ценность альтернативных вариантов и осуществлять выбор релевантного целям пользователя варианта. Рассмотренный пример сравнения вариантов образов демонстрирует работоспособность предложенного подхода и возможность его применения для выбора стилистических решений.

Библиографический список

1. David Zyla Color Your Style / Penguin Books, 2011.
2. Баранов В.Г., Седаков Д.В., Милов В.Р. Способ управления конфигурацией комплекса информационных систем // Информационно-измерительные и управляющие системы. 2017. № 8. С. 26-33.
3. Yegorov Y.S., Milov V.R., Kvasov A.S., Sorokoumova S.N., Suvorova O.V., Formalization of software requirements for information systems using fuzzy logic // IOP Conf. Series: Journal of Physics: Conf. Series 1015 (2018) 042062 doi:10.1088/1742-6596/1015/4/042062.

Y.S. Yegorov, E.A. Kamyshan

**AN APPROACH TO DEVELOPING AN INTELLIGENT ADVISORY SYSTEM
FOR SUPPORTING STYLISTIC SOLUTIONS USING FUZZY LOGIC**

Nizhny Novgorod State Technical University n.a. R. E. Alekseev

The paper considers an approach to describing and choosing stylistic solutions for the selection of the most relevant clothing options and their combinations. The approach allows you to formalize the requirements using the apparatus of fuzzy logic and to define particular objective functions to support the adoption of stylistic decisions.

Keywords: decision support, information system, design, stylistic decisions, fuzzy logic.

РАЗРАБОТКА ПРОЦЕДУР ИНТЕЛЛЕКТУАЛЬНОЙ ОЦЕНКИ ЗНАНИЙ НА ОСНОВЕ СЕМАНТИЧЕСКОЙ ОБРАБОТКИ ОТВЕТОВ ПОЛЬЗОВАТЕЛЕЙ НА ЕСТЕСТВЕННОМ ЯЗЫКЕ

Нижегородский государственный технический университет им. Р.Е. Алексеева

В работе предлагается методика интеллектуальной оценки знаний на основе ответов пользователей на открытые вопросы. Рассмотрены процедуры предварительной обработки, векторизации текста и классификации векторного представления. Представлены результаты обучения и тестирования алгоритмов.

Ключевые слова: обработка естественного языка, проверка знаний, система тестирования, семантический анализ текста.

Процесс обучения в любой образовательной организации подразумевает проверку качества усвоения полученных знаний учащимися с целью установления обратной связи с педагогом для последующей корректировки образовательного процесса.

Тестирование является эффективным видом проверки знаний студентов и играет важную роль при оценке результатов обучения. Наиболее продуктивным методом можно назвать тестирование с помощью открытых вопросов. Такой метод требует от студента умения корректно и четко формулировать свои мысли и позволяет давать развернутые ответы на естественном языке, при этом практически полностью исключая возможность случайного выбора правильного ответа [1].

Ответы на такие вопросы педагоги обычно проверяют вручную, что является достаточно трудоемкой задачей. Автоматическое оценивание призвано решить эту проблему, а также улучшить другие аспекты обучения за счет автоматизации процесса выставления оценок.

Автоматическая проверка знаний может обеспечить снижение временных затрат педагогов на проверку ответов, получение быстрой обратной связи и улучшение объективности оценивания, исключив ошибки проверки из-за усталости или невнимательности [2-3].

Таким образом, исследование направлено на разработку технологии, обеспечивающей автоматическое оценивание ответов учащихся на открытые вопросы, что позволит реализовать интеллектуальную высокоавтоматизированную обработку запросов пользователей на естественном языке применительно к системам тестирования знаний с открытыми вопросами.

Методика оценивания ответов на открытые вопросы

Предлагаемая методика решения задачи оценивания ответов на открытые вопросы включает этапы предварительной обработки текста вопросов и ответов, представления текста в виде числовых векторов и классификации полученных векторов. Основные этапы представлены на рис. 1.

Алгоритмы классификации не могут напрямую работать с сырым текстом, поэтому необходимо конвертировать текст в наборы чисел – векторы. В рамках работы для векторизации текста рассматривался «классический» метод мешка слов.

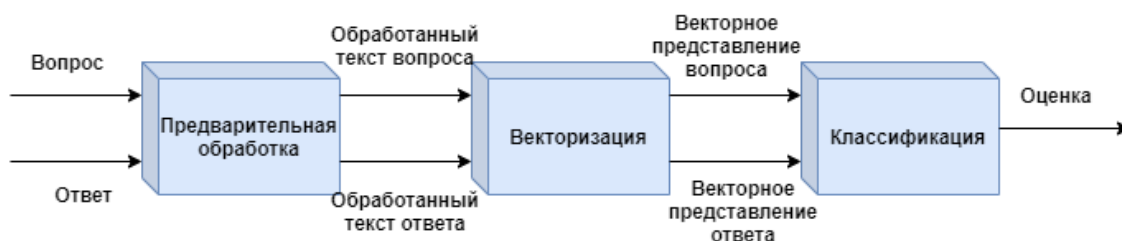


Рис. 1. Этапы оценивания ответов на открытые вопросы

Векторное представление вопроса и ответа тестируемого поступает на вход процедуры классификации для автоматизированного формирования оценки за ответ. Для реализации процедуры классификации рассматривались следующие методы:

1. Метод опорных векторов
2. Случайный лес
3. Метод k-ближайших соседей
4. Искусственная нейронная сеть

Такие модели на выходе формируют числовое значение, характеризующее принадлежность входных вопроса и ответа к классу оценки.

Использование указанного векторизатора и методов классификации позволит сравнить полученные результаты и выявить лучшие методы, а также оценить применимость предложенной методики для решения задачи оценивания ответов на открытые вопросы.

Набор данных

Для обучения и тестирования рассматриваемых процедур использовался англоязычный набор ответов на открытые вопросы SciEntsBank (англ. The Science Entailments corpus), который был собран для конкурса SemEval-2013 Task 7 [4]. Датасет состоит из вопросов, ответов учащихся, справочных ответов и оценок, определенных в порядковой шкале. Вопросы относятся к разным предметным областям, таким как физика, науки о жизни, земле, космосе и т.д.

Каждому ответу присвоена одна из следующих оценок:

- «correct» - полностью правильный или аналогичный справочному ответ;
- «partially_correct_incomplete» - частично правильный ответ;
- «contradictory» - противоречащий справочному ответ;
- «irrelevant» - неуместный ответ в рамках предметной области;
- «non_domain» - ответ, не относящийся к предметной области и демонстрирующий отсутствие знаний (например, «I don't know»).

Фрагмент подготовленной выборки представлен в таблице 1.

На основе данных SciEntsBank была сформирована выборка, содержащая 196 уникальных вопросов по 15 предметным областям с 11000 ответами, включая 196 справочных ответов (по одному на вопрос) и 10804 ответов студентов. Количество вопросов составляет от 3 до 25 для различных предметных областей. Для большинства предметных областей среднее количество ответов на каждый из вопросов составляет 40, для нескольких областей имеется около 100 ответов.

Для обучения и тестирования процедуры оценивания ответов на открытые вопросы набор данных SciEntsBank разделялся на обучающую и тестовую выборки. В выборках присутствовали одинаковые вопросы по одинаковым предметным областям, однако тестовая выборка содержала новые ответы для каждого вопроса, которые не входили в обучающий набор. При этом при разбиении данных применялся метод стратификации, позволяющий выбрать объекты таким образом, чтобы сохранялось соотношение меток классов («correct», «partially_correct_incomplete», «contradictory», «irrelevant», «non_domain») в обучающем и тестовом наборах.

Фрагмент таблицы с данными из набора SciEntsBank

module	id question	question	id_ answer	answer	grade
EM	EM_13	What is the difference between a rock and a mineral?	EM.13.521.1	Material made up of different ingredients called mineral. And rocks can be squished up and mineral can they are squish up all very squished up.	contradictory
EM	EM_13	What is the difference between a rock and a mineral?	EM.13.536.1	A rock is something from the ground and it has its own property. A mineral has its own properties also. But it is different from a rock because minerals have their own name.	irrelevant

Предварительная обработка и векторизация текста

При обработке текста вопросов и ответов удалялись все знаки препинания. Затем текст приводился к нижнему регистру и исключались английские стоп-слова.

Для векторизации текста вопросов и ответов рассматривались варианты с разными значениями n-грамм (1, 1 и 2, 2) для поиска наилучшего решения, а также 2 векторизатора – один для вопросов, другой для ответов. Полученные вектора признаков конкатенировались в общий вектор, подававшийся на вход классификатора. Его соответствовала количеству уникальных слов и/или фраз в тексте в зависимости от значения n-грамм.

Значения векторов признаков зависят от метода векторизации:

1. Векторизация по количеству (соответствует количеству вхождения уникальных слов и/или фраз в вопрос и ответ).
2. Векторизация по мере tf-idf (соответствует значению меры tf-idf для каждого уникального слова и/или фразы).

Обучение и тестирование классификаторов

Векторное представление вопросов и ответов использовалось для обучения и тестирования ряда алгоритмов классификации применительно к задаче оценки ответов пользователей на открыты вопросы.

Оценка качества алгоритмов проводилась путем 5-кратной перекрестной проверки – 80% данных в обучающей выборке и 20% данных в тестовой выборке. Метрики правильности (accuracy), точности (precision) и полноты (recall) рассчитывались для каждой итерации 5-кратной перекрестной проверки. По полученным значениям метрик рассчитывались результирующие значения как среднее арифметическое для эксперимента.

Согласно полученным результатам для простых классификаторов (метода опорных векторов, случайного леса и k-ближайших соседей), правильность (accuracy) классификации достигает 89-91 % на обучающей выборке и 63-65 % – на тестовой выборке. Векторизация по количеству вхождений слов и по мере tf-idf дает близкие результаты.

Так как обучение искусственной нейронной сети является достаточно ресурсоемкой задачей, то для этого алгоритма был выбран оптимальный метод векторизации – векторизация по количеству вхождений слов с использованием юниграмм (N=1).

Метод Dropout и L2 регуляризация к ядру слоя применялись для стабилизации процесса обучения. В качестве функции активации нейронов использовалась функция ReLU, в последнем слое – функция Softmax. Обучение длилось 200 эпох со скоростью обучения 0,00001, используемая функция потерь – категориальная кросс-энтропия.

Разработанная модель нейронной сети является стабильной, что подтверждается минимальной разностью значений метрик и функции потерь для разных итераций кросс-валидации. В результате обучения было достигнуто среднее значение правильности 98% на обучающей выборке и наилучший результат **70,49%** на тестовой выборке (что превосходит результаты предыдущих классификаторов).

Усредненные значения метрики правильности по всем итерациям кросс-валидации для рассмотренных классификаторов – метода опорных векторов, случайного леса, k-ближайших соседей и искусственной нейронной сети – представлены в таблице 2.

Таблица 2

Результаты обучения и тестирования простых ML классификаторов

Метод классификации	Метод векторизации	Значение N-грамм	Accuracy	
			Обучающая выборка	Тестовая выборка
Метод опорных векторов	Количество вхождений	1	0,8048	0,6388
		1 и 2	0,9688	0,6434
		2	0,9736	0,6172
	Мера tf-idf	1	0,7819	0,6578
		1 и 2	0,9212	0,6672
		2	0,9352	0,6434
Случайный лес (200 деревьев)	Количество вхождений	1	0,9981	0,7023
		1 и 2	0,9989	0,6896
		2	0,9946	0,6443
	Мера tf-idf	1	0,9981	0,6993
		1 и 2	0,9989	0,6866
		2	0,9946	0,6494
K – ближайших соседей (k = 3)	Количество вхождений	1	0,7924	0,6344
		1 и 2	0,7699	0,6054
		2	0,7168	0,5461
	Мера tf-idf	1	0,8101	0,6527
		1 и 2	0,8033	0,6442
		2	0,7588	0,5883
Искусственная нейронная сеть	Количество вхождений	1	0,9855	0,7049

Таким образом, согласно представленным значениям метрик для различных подходов, предложенная методика показывает достаточно перспективные результаты, которые в дальнейшем могут быть улучшены путем применения современных нейросетевых алгоритмов обработки естественного языка (BERT, GPT-3 и др.).

Библиографический список

1. Кожевников В.А., О.Ю. Сабинин Система автоматической проверки ответов на открытые вопросы на русском языке // Научно-технические ведомости СПбГПУ. Информатика, телекоммуникации и управление. 2018. №11:3. С.57–72.
2. Милов В.Р., М.С. Дубов, Н.А. Калинина, А.А. Салтыкова Интеллектуализация тестирования с открытыми вопросами на основе определения семантической близости фраз // Интеллектуальные системы в науке и технике. Искусственный интеллект в решении актуальных социальных и экономических проблем XXI века. Сборник статей по материалам Международной конференции и Шестой всероссийской научно-практической конференции. 2020. С.112-115.
3. Дубов М.С., В.Р. Милов, Н.А. Калинина, А.А. Салтыкова Процедура автоматизированного оценивания свободных ответов обучающихся // Нейрокомпьютеры и их применение. XVIII Всероссийская научная конференция. 2020. С. 88-89.

4. The Joint Student Response Analysis and 8th Recognizing Textual Entailment Challenge [Электронный ресурс]: - Режим доступа: <https://www.cs.york.ac.uk/semEval-2013/task7/index.html>.

A.Y. Kerbeneva

DEVELOPMENT OF PROCEDURES FOR INTELLECTUAL KNOWLEDGE ASSESSMENT BASED ON SEMANTIC PROCESSING OF USERS' ANSWERS IN NATURAL LANGUAGE

Nizhny Novgorod State Technical University n.a. R.E. Alekseev

The paper proposes a methodology for intellectual assessment of knowledge based on users' answers to open questions. The procedures for preprocessing, text vectorization and classification of vector representation are considered. The results of training and testing algorithms are presented.

Keywords: natural language processing, knowledge testing, testing system, semantic text analysis.

ПОДХОД К РАЗРАБОТКЕ СИСТЕМЫ РАСПОЗНАВАНИЯ ЭЛЕМЕНТОВ ДОРОЖНОЙ ИНФРАСТРУКТУРЫ С ИСПОЛЬЗОВАНИЕМ ТЕХНОЛОГИИ КОМПЬЮТЕРНОГО ЗРЕНИЯ

Нижегородский государственный технический университет им. Р.Е. Алексеева

В работе выполнены обзор и выбор методов обработки видеоизображения при реализации информационной системы, позволяющей осуществлять захват изображения с видеокамеры в салоне автомобиля, распознавать и сопровождать элементы дорожной инфраструктуры с целью информирования водителей о состоянии дорожного полотна.

Ключевые слова: информационная система, обработка данных, нейронная сеть, дорожная инфраструктура.

Качество дорожного покрытия – одна из самых острых проблем не только России, но и всего мира. Практически каждый день водители используют личный транспорт, а пассажиры общественный. Со временем дорожное покрытие изнашивается и требуется его замена.

Согласно публикациям экспертно-аналитического центра Российской академии народного хозяйства и государственной службы при Президенте Российской Федерации, порядка 40% ДТП происходит на участках дорог «не соответствующих режимам движения» и плохих по качеству трассах, при этом в России ремонта требуют около 70% федеральных дорог РФ, в регионах дорожное покрытие не меняется иногда десятилетиями [1].

Современные технологии машинного зрения и методы интеллектуальной обработки данных открывают новые возможности обеспечения безопасности дорожного движения. На данный момент машинное зрение уже активно применяется, например, при определении в дорожном потоке номера автомобиля, его марки, фактов нарушения правил дорожного движения, а также при фиксации дорожной разметки.

Каждый год ведётся ремонт автомобильных дорог, улучшение дорожной сети.

Существуют следующие виды работ, выполняемые по их содержанию:

- Текущий ремонт – систематически плано-предупредительные работы по предупреждению и исправлению мелких повреждений дороги и её сооружений, проводимые в течение всего года на всем протяжении дороги.
- Капитальный ремонт – периодические комплексные работы, выполняемые один раз в несколько лет, направленные на полное восстановление основных эксплуатационных качеств дороги и дорожных сооружений.
- Реконструкция автомобильной дороги – совокупность работ и мероприятий, направленных на изменение и повышение основных технико-экономических показателей и потребительских свойств существующей автомобильной дороги или отдельных её участков и дорожных сооружений.

При проведении ремонтных работ используются ограждения, которые ограничивают поток движения, а водители при этом часто пытаются найти пути объезда, чтобы сэкономить время в пути.

Сложившаяся обстановка предопределяет необходимость совершенствования процесса информирования водителей и пассажиров, что становится возможным за счет создания системы распознавания элементов дорожной инфраструктуры с использованием технологии компьютерного зрения.

Разрабатываемая информационная система на основе нейросетевых процедур должна своевременно информировать участников дорожного движения и соответствующих службы о состоянии дорожной инфраструктуры и дорожного полотна.

Примером работы информационной системы может служить обнаружение и сопровождение такого элемента дорожной инфраструктуры, как дорожный конус (рис. 1).

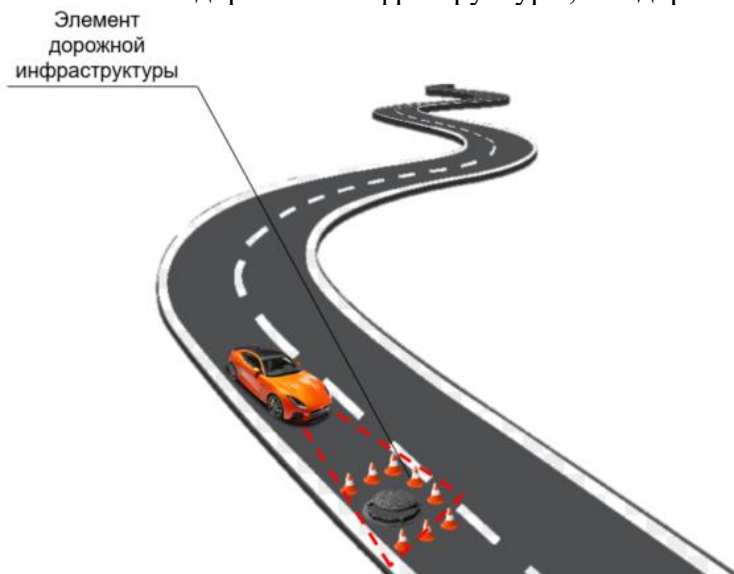


Рис. 1. Пример работы системы распознавания элементов дорожной инфраструктуры

Информационная система должна включать в свой состав следующие модули:

1. Модуль ввода данных – осуществляет взаимодействие с видеочамерой.
2. Модуль обработки данных – осуществляет внутреннюю логику обработки изображений, распознавания и сопровождения объектов.
3. Модуль вывода данных – осуществляет вывод информации на графический интерфейс пользователя.

Идентификация дорожной инфраструктуры может осуществляться несколькими последовательно применяемыми алгоритмами.

Исходная видеозапись обрабатывается путем анализа каждого кадра. На основе захваченного кадра $x(n)$ и обработанного ранее кадра $y(n-1)$ происходит обнаружение элементов дорожной инфраструктуры. Обнаружение позволяет определить местоположение объекта в кадре.

Результаты применяются для последующего сопровождения. Обработанный кадр с выделенными в рамки элементами $y(n)$ сохраняется в компоненте памяти для анализа следующего поступающего кадра [2].

Структура разрабатываемой информационной системы представлена на рис. 2.

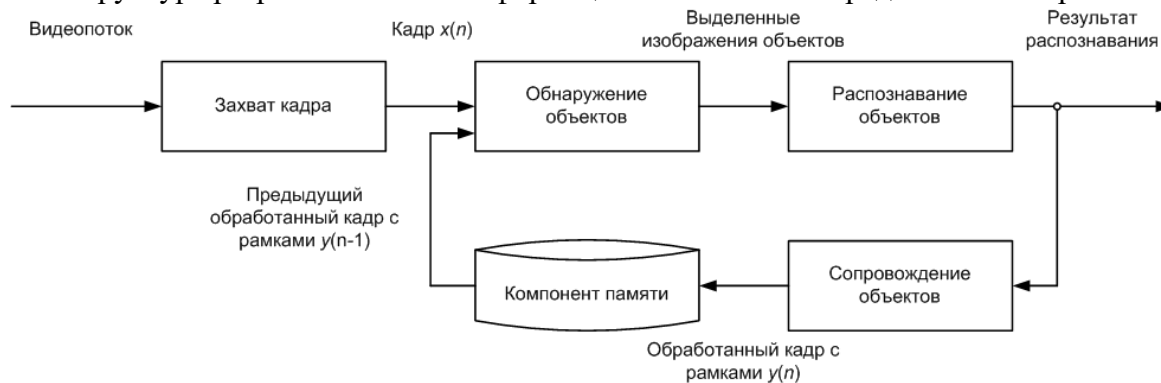


Рис. 2. Структура разрабатываемой информационной системы

В рамках рассматриваемого подхода к разработке системы распознавания элементов дорожной инфраструктуры могут использоваться следующие методы.

Метод обнаружения объектов

Метод Виолы-Джонса — алгоритм обнаружения объектов на изображениях в режиме реального времени. Схема каскадной структуры представлена на рис. 3.

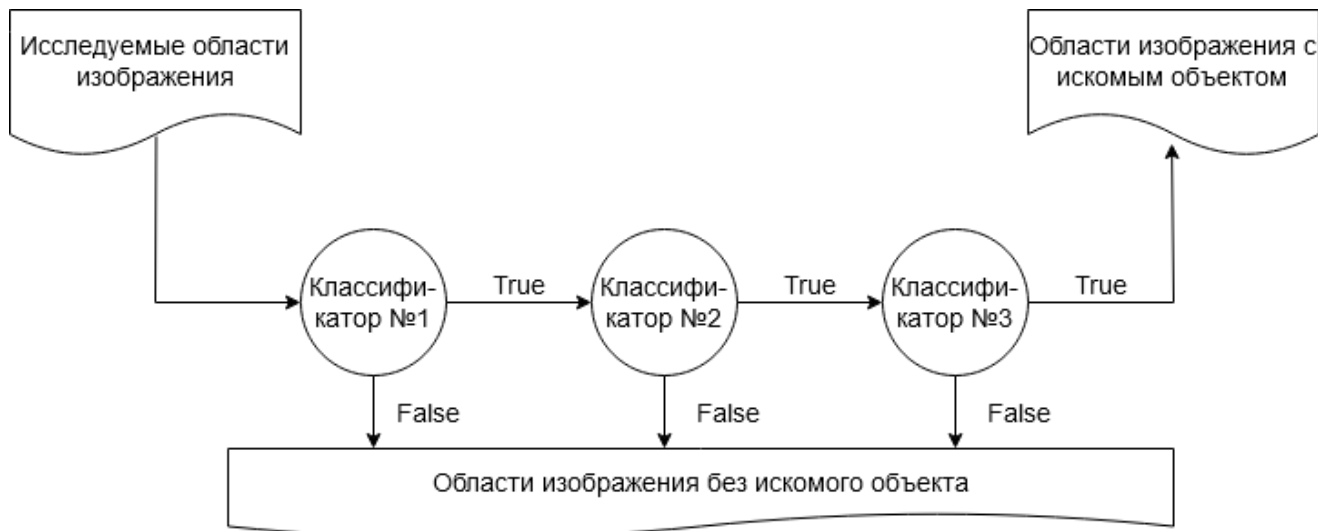


Рис. 3. Схематическое описание каскадной структуры

Метод основан на следующих принципах [3]:

- изображение представляется в интегральной форме и обрабатывается скользящим окном, что позволяет быстро вычислять искомый объект;
- признаковое описание объекта строится с помощью примитивов Хаара;
- признаковое описание используется для построения бинарных классификаторов;
- применяется алгоритм адаптивного усиления (англ. Adaptive Boosting, AdaBoost), позволяющий выбрать наиболее эффективные признаки и усилить слабые бинарные классификаторы, настроенные на каждый отобранный признак;
- сильные классификаторы объединяются в каскады для быстрого обнаружения регионов, потенциально содержащих целевой объект.

Каскад сверточных нейронных сетей (англ. Multi-Task Cascaded Convolutional Neural Networks, MTCNN) – несколько сверточных нейронных сетей, каждая из которых нацелена на решение отдельной конкретной подзадачи. Такие сети обучаются по принципу многозадачного обучения (англ. Multi-Task Learning, MTL).

Многозадачное обучение позволяет решать несколько задач, при этом используя общие черты и различия между ними, изучая их совместно. Такой подход приводит к повышению эффективности обучения и точности прогнозирования моделей, специфичных для конкретной задачи.

Методы сопровождения объектов

Метод сопровождения особых точек (англ. Point Tracking). Можно выделить два этапа работы такого алгоритма. На этапе детектирования определяются особенности изображения на первом кадре. Для следующих кадров определяется положение особенности (этап отслеживания). В качестве особых точек можно рассматривать углы, ребра и другие части изображения.

Метод сопровождения компонент (англ. Kernel Tracking) [4]. Под компонентой подразумевается форма объекта, которая в простейшем случае может быть представлена прямоугольным или овальным шаблоном. Метод применяется в тех случаях, когда движение объекта представляет собой обычное смещение, поворот или аффинное преобразование.

Метод сопровождения с помощью сопоставления с расширенной моделью появления объекта (англ. Incremental Visual Tracking) [4] применяется для долгосрочного отслеживания. Идея заключается в сохранении расширенной модели поведения цели, полученной с предыдущих кадров. Собственные признаки целевого объекта вычисляются с помощью инкрементного метода главных компонент над шаблоном значения интенсивности объекта.

Метод сопровождения с помощью сопоставления с ограничениями направлен на то, чтобы отслеживать цели, для которых форма объекта сильно изменяется со временем, с помощью разреженной оптимизации по парам фрагментов изображения (сопровождение методом Монте-Карло) [4]. С учётом представленного местоположения объекта на первом кадре, цель моделируется путем выборки фиксированного количества целевых фрагментов, которые описываются краевыми элементами и гистограммами цвета.

Метод сопровождения силуэта (англ. Silhouette Tracking) [4]. Силуэт может быть представлен в виде контура или набора связанных простых геометрических примитивов. Задача алгоритма заключается в том, чтобы определить положение объекта на каждом кадре, используя модель его силуэта, построенную на предыдущих кадрах.

Метод сопровождения с помощью различающей классификации заключается в том, чтобы построить модель по отличию целевого переднего плана от фона [4]. Сопровождение путем обнаружения создает классификатор, чтобы отличать целевые пиксели от фоновых пикселей, и обновляет классификатор новыми входящими в него образцами.

Метод сопровождения с использованием корреляционных фильтров [4]. Корреляционные фильтры могут сопровождать объекты в сложных условиях при наличии вращений, частичных исчезновений объекта и других отвлекающих факторов быстрее других подходов. Внешний вид цели моделируется адаптивными корреляционными фильтрами, причем сопровождение осуществляется посредством операции свертки.

Выбор методов для решения поставленной задачи

В качестве подхода к решению задачи предполагается использовать [5]:

- Метод обнаружения объектов – каскады сверточных нейронных сетей.
- Метод вычисления признаков – сиамские сверточные нейронные сети.

Такой выбор обусловлен тем, что на данный момент сверточные нейронные сети и их модификации считаются лучшими по точности и скорости алгоритмами нахождения объектов на сцене и дальнейшей их классификации. Пример модели сверточной нейронной сети представлен на рисунке 4 [6].

К достоинствам сверточных нейронных сетей перед традиционными вычислительными системами можно отнести следующее [7]:

1. Используя способность обучения на множестве примеров, нейронная сеть способна решать задачи, в которых неизвестны закономерности развития ситуации и зависимости между входными и выходными данными.
2. Нейронные сети обладают способностью адаптироваться к изменениям окружающей среды.
3. Удобное распараллеливание вычислений, а, следовательно, возможность реализации алгоритмов работы и обучения сети на графических процессорах.
4. Относительная устойчивость к повороту и сдвигу распознаваемого изображения.

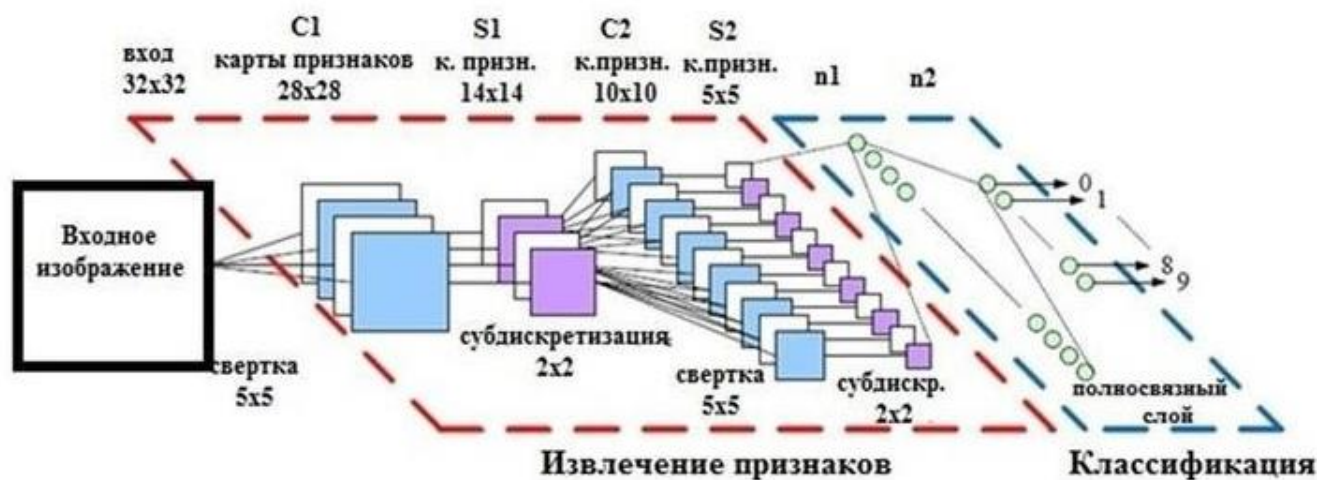


Рис. 4. Пример модели сверточной нейронной сети

В разрабатываемой информационной системе также предполагается использовать мини-максную нормализацию векторов признаков и осуществлять их сравнение путем расчета Евклидова расстояния. Такой метод является одним из традиционных при оценке схожести векторов и не требует обучения классификаторов.

Выбор метода сопровождения объектов будет проведен по результатам тестирования существующих алгоритмов.

Подход, содержащий совокупность указанных методов, позволит решить следующие задачи, связанные с разработкой информационной системы:

1. Разработка информационной модели информационной системы.
2. Разработка алгоритма работы и процедуры обработки данных.
3. Создание экспериментального образца системы распознавания элементов дорожной инфраструктуры.

Библиографический список

1. Статистика ДТП [Электронный ресурс]. –URL: <https://iz.ru/news/601872>. Дата обращения: 25.01.2021.
2. Краткий курс машинного обучения [Электронный ресурс]. - URL: <https://netpeak.net/ru/blog/kratkiy-kurs-mashinnogo-obucheniya-ili-kak-sozdat-neuronnyu-set-dlya-resheniya-zadachi-po-skoringu/>. Дата обращения: 25.01.2021.
3. Выделение объектов на изображении по методу Виоллы-Джонса [Электронный ресурс] URL: <https://api-2d3d-cad.com/viola-jones-method/>. Дата обращения: 25.01.2021.
4. Каллан, Р. Основные концепции нейронных сетей/ Роберт Каллан – М.: Изд. дом. «Вильямс», 2001. – 287 с.
5. Форсайт Д. Компьютерное зрение. Современный подход/ Д. Форсайт – М.: Изд. дом. «Вильямс», 2004. – 928 с.
6. Видеоаналитика. Термины, сферы применения, технологии Video Content Analysis [Электронный ресурс]. URL: [https://www.tadviser.ru/index.php/Статья:Видеоаналитика_\(термины,_сферы_применения,_технологии\)](https://www.tadviser.ru/index.php/Статья:Видеоаналитика_(термины,_сферы_применения,_технологии)). Дата обращения: 25.01.2021.
7. Компьютерное зрение. Задачи, области применения, перспективы [Электронный ресурс] URL: <https://vc.ru/ml/166105-kompyuternoe-zrenie-zadachi-oblasti-primeneniya-perspektivy>. Дата обращения: 25.01.2021.

Y.A. Klimin, Y.S. Yegorov

**AN APPROACH TO THE DEVELOPMENT OF A SYSTEM FOR RECONGIZING
ROAD INFRASTRUCTURE ELEMENTS USING COMPUTER VISION TECHNOLOGY**

Nizhny Novgorod State Technical University n.a. R.E. Alekseev

The paper considers an approach that allows capturing images from a video camera in the car interior, recognizing and accompanying elements of road infrastructure in order to inform drivers about the state of the roadway.

Keywords: information system, data processing, neural network, road infrastructure.

РАЗРАБОТКА СИСТЕМЫ УПРАВЛЕНИЯ БИЗНЕС-ПРОЦЕССАМИ ПРОДАЖ С ПРИМЕНЕНИЕМ ИГРОВЫХ МЕТОДОВ ВОВЛЕЧЕНИЯ ПОЛЬЗОВАТЕЛЕЙ

Нижегородский государственный технический университет им. Р.Е. Алексеева

В работе описывается процесс разработки системы управления бизнес-процессами продаж с применением игровых методов вовлечения пользователей. Основываясь на обзоре существующих программных решений, доступных на момент написания происходит анализ и выбор методов разработки, а также выбор средств реализации как программно-аппаратной части, так и пользовательского интерфейса. Также выбирается методология и разрабатывается схема базы данных. Путем эмпирического исследования разработаны алгоритмы реализации игровых заданий для пользователей и алгоритм анализа статистики выполнения заданий.

Ключевые слова: игрофикация, геймификация, бизнес-процесс, информационная система, маркетинг, социальная сеть.

Современные компании значительно увеличивают бюджет на маркетинг и рекламные кампании для привлечения, вовлечения и удержания потенциальных клиентов [1]. Конечная задача обычно заключается в том, чтобы клиенты приобрели товар или воспользовались рекламируемой услугой. Из-за того, что конкуренция на рынке постоянно растет, а потенциальные клиенты всё хуже вовлекаются в процессы, которые результативно работали еще несколько лет назад – падают реальные доходы компаний. Модернизация бизнес-процессов для оптимизации расходов компании на привлечение потенциального клиента, а также модернизация способов удержания вовлеченности клиента для дальнейшей продажи продукта или услуги – это основная задача многих бизнесов сегодня [2]. Данную задачу помогает решить введение игровых механик в неигровые процессы, используя новейшие разработки из психологии и индустрии компьютерных игр.

Основная задача геймификации – создать условия для изменения поведения потенциальных клиентов [3]. Геймификация позволяет избежать негативных реакций на рекламу и установить практически не рекламную работу с клиентом, улучшенную, по сравнению с теми, кто её не использует. Бизнес становится более активным, больше взаимодействует с аккаунтами в социальных сетях, а значит, больше покрытия.

Пользователи, вовлеченные в игру, охотнее делятся информацией с друзьями. В результате геймификация может предоставить более дешевые приложения от потенциально заинтересованных клиентов по сравнению с другими форматами рекламы [4].

Геймификация позволяет:

- создать необходимые ассоциативные ссылки;
- увеличить количество взаимодействий с контентом;
- привлекать новых клиентов и удерживать старых;
- вызвать потребность в продукте;
- увеличить количество продаж, в том числе повторных, и средний чек;
- вызывать положительные эмоции от взаимодействия с брендом;
- повысить доверие пользователей;
- обучать сложным процессам и продуктам [5].

Для того, чтобы применять геймификацию в реальных продуктах, необходимо решить проблему внедрения её элементов в бизнес-процессы.

Такая задача решается преимущественно внедрением системы управления бизнес-процессами.

В структуре разработанной системы можно выделить шесть составных модулей: базу данных, модуль обработки запросов, модуль формирования заданий, модуль анализа, модуль связи по API с социальной сетью (реализовано на примере социальной сети «ВКонтакте»), модуль анализа выполнения заданий, модуль формирования ссылок на оплату заказов. На рис. 1 представлена контекстная диаграмма.

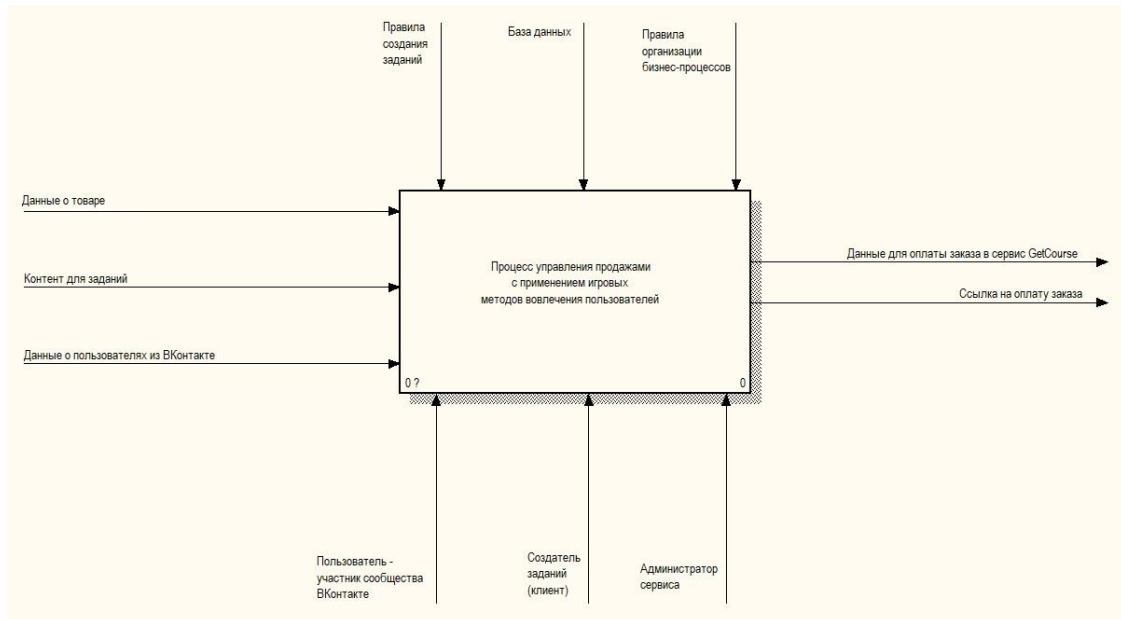


Рис. 1. Контекстная диаграмма

На рис. 2 представлена диаграмма декомпозиции.

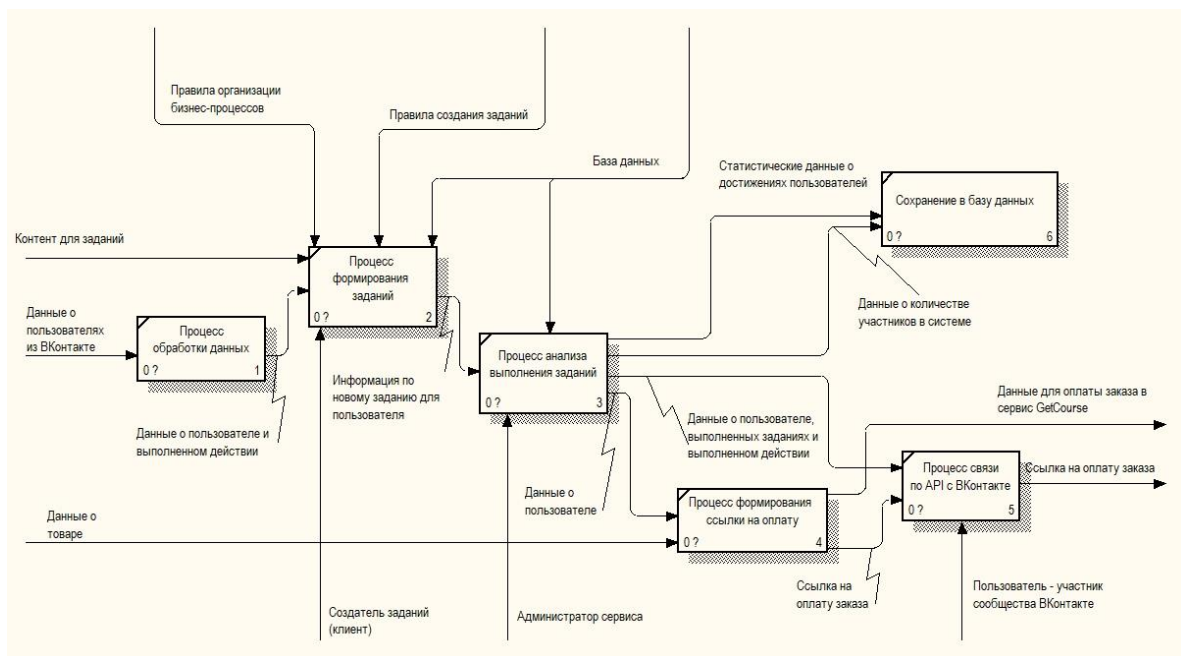


Рис. 2. Диаграмма декомпозиции

Модуль сохранения в базу данных – это модуль, который отвечает за сбор информации о пользователях и клиентах системы, основанный на взаимодействии с таблицами MySQL.

Модуль обработки данных – это специальный модуль, который отвечает за процесс обработки всех входящих запросов из социальной сети ВКонтакте. Все запросы веб-сервер направляет в единую точку входа – файл index.php в директории vk.

Модуль связи по API с ВКонтакте – модуль, отвечающий за выполнение обработки и запросов к публичному API социальной сети ВКонтакте.

Модуль формирования заданий – модуль, преобразующий входные данные из пользовательского интерфейса клиента в код, который записывается в выделенные директории под хранение заданий.

Модуль анализа заданий – модуль, который взаимодействует с базой данных, получает данные по пользователю, отправившему запрос, а затем сравнивает найденные значения с той структурой обработки и анализа, которая задается через графический интерфейс.

Модуль формирования ссылки на оплату – модуль, который отвечает за взаимодействие с внешними системами. Он посылает данные в указанном формате на url нужного сервиса и ожидает ответ, содержащий ссылку на оплату и данные о товаре, которые впоследствии будут отправлены пользователю для выполнения покупки.

Правила организации бизнес-процессов, правила создания заданий и контент для заданий формируются на стороне клиента и зависят исключительно от его бизнес-задач. Чаще всего такие правила прописываются бизнес-аналитиками и затем реализуются в системе.

Алгоритм реализации клиентом игровых заданий для пользователей

На рис. 3 представлена блок-схема алгоритма реализации клиентом игровых заданий для пользователей.

Блок-схема алгоритма реализации клиентом игровых заданий для пользователей состоит из следующих блоков:

- заполнение формы создания заданий – клиент указывает название для будущего алгоритма заданий и указывает тип получаемого запроса с API ВКонтакте, на который будет включаться этот алгоритм;
- создание директории для алгоритмов и создание файлов с алгоритмами и переменными – система создает новую директорию, в которой будет храниться информация о цепочке действий и условий, которые будут выполняться при запуске самого алгоритма;
- клиент создает последовательность условий и действий – система открывает страницу с возможностью добавления блоков условий и действий. Клиент создает цепочку из готовых блоков вида: условия вхождения и проверки, затем действия, которые будут выполняться при прохождении этих условий;
- генерация кода для условий в файл задания – в коде файлов заданий формируются условные конструкции с пометкой номеров блоков, которые им соответствуют. Условные конструкции имеют строгий тип: если не выполнится одно из них, то второе уже не дойдет до проверки. Сами методы проверки условий изначально расположены в директориях системы и запрашиваются по названиям файлов. Переменные для алгоритмов формируются из атрибутов методов и перезаписываются с указанием блока условия, которому они принадлежат;
- генерация кода для действий в файл задания – в коде файлов заданий после условий создаются новые действия из предустановленных методов. Переменные для действий формируются таким же образом, как и для условий с отличием в том, что нумерация блоков создается исключительно для блоков действий, исключая блоки условий.

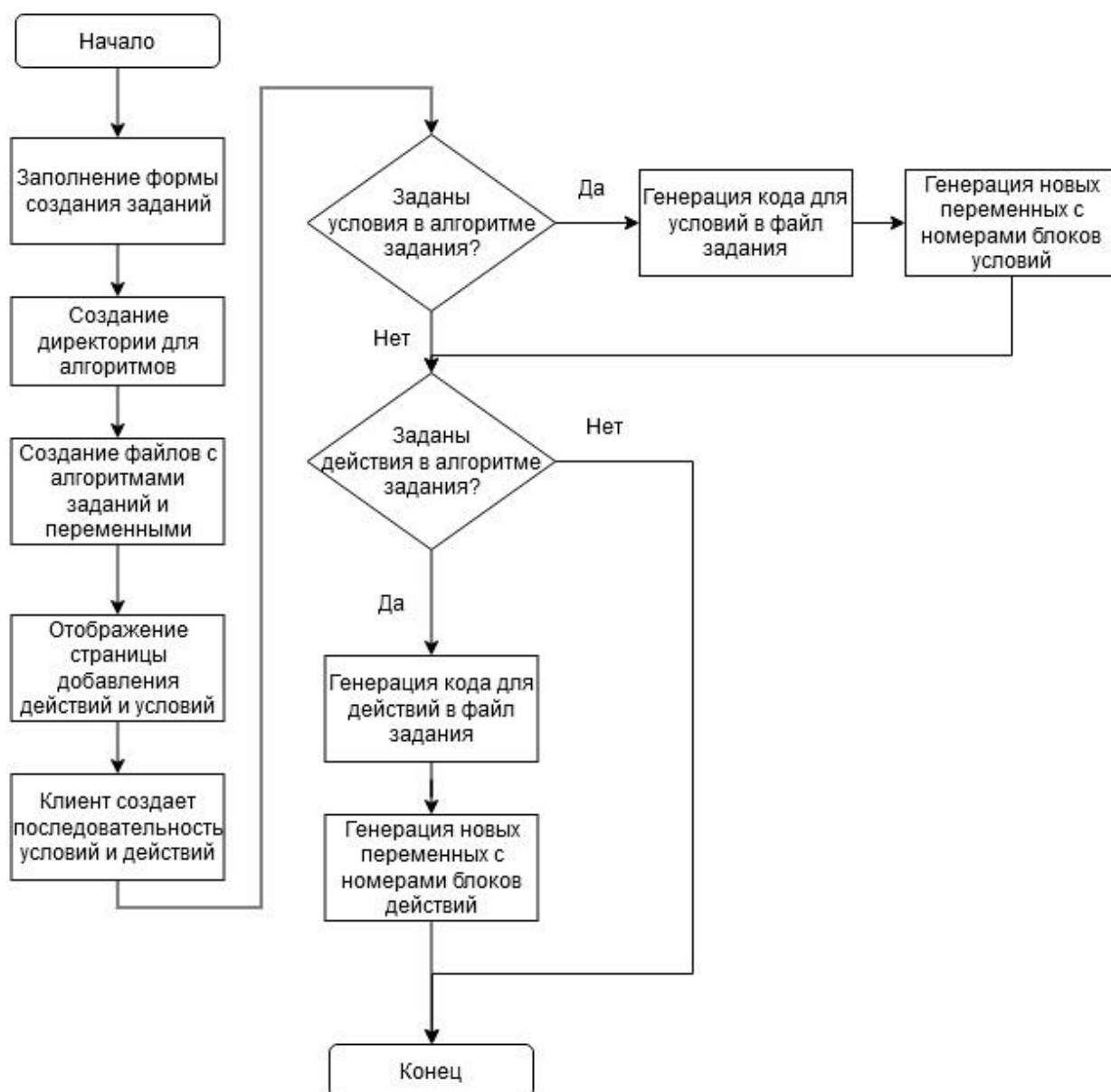


Рис. 3. Блок-схема алгоритма реализации клиентом игровых заданий для пользователей

Пользователь системы через графический интерфейс задает информацию о том, за какое действие клиента будет назначаться награда, какая будет награда и какое количество. После сохранения настроек, система запоминает этот алгоритм в текстовый файл с PHP кодом.

При следующем запросе от API ВКонтакте система производит проверку по всем созданным заданиям пользователя, которые точно подходят под заданный тип условия выполнения цепочки задания; затем анализирует алгоритм бота (задания) на выполнение блоков условий. Если хоть какое-то из условий, заданное клиентом, не выполняется, то выполнение алгоритма прекращается. В ином случае происходит дальнейшее выполнение всех методов действий, которые задал клиент. Действия могут быть связаны как с внутренней работой системы: взаимодействие с базой данных, установка новых значений в боте; так и с внешними системами: передача запросов во ВКонтакте по API, запросы на сторонние сервисы, POST и GET запросы [6].

К примеру, клиент системы установил, что при событии «Новое сообщение» с текстом сообщения равным «Балл», клиенту будет начислять дополнительно плюс 1 балл к его общему балансу. Когда клиент напишет в диалог сообществу пользователя слово

«Балл», то система получит эти данные, обработает их и в базе данных запишет новое значение для пользователя по такому алгоритму: если клиент с таким vk id уже существует и имеет запись баллов, то к его счету прибавится 1 балл; если такого клиента ещё не существует в базе данных, то система создаст новую запись со значением на балансе, равным 1 балл. Балл в данном примере может являться записью о выполнении задания или же достижением, которое увидит пользователь.

Алгоритм анализа статистики выполнения заданий

Анализ статистики основан на работе со значениями транзакций баллов, сохраняемых в базе данных. В основе алгоритма заложено, что участник сообщества, выполняя задание пользователя системы в его сообществе социальной сети ВКонтакте, должен каким-либо образом идентифицировать выполнение этого задания. Например, написав ключевое (уникальное) слово в тексте ответа, либо же написав свой ответ в определенном месте, указанном клиентом системы.

Такие условия для сбора данных о выполнении заданий задаются клиентом через графический интерфейс системы. По этим условиям система сможет обрабатывать сохранять уникальные данные в соответствующие ячейки сбора информации, когда будет поступать запрос из социальной сети ВКонтакте с соответствующим типом.

Анализ выполнения заданий заключается в просмотре и формировании статистики по заданным клиентом значениям баллов в системе.

На рис. 4 представлена блок-схема алгоритма анализа выполнения заданий пользователем.

Описание блок-схемы:

- поступление callback запроса из ВКонтакте – данные о действиях участников сообщества пользователя поступают на обработчик системы, в виде json строки с параметрами самого действия, информации об участнике, который это действие совершил и дополнительно в запросе прикладываются данные с id сообщества и его секретным ключом;
- проверка в базе данных наличия сообщества, из которого пришёл запрос – система проверяет наличие в базе данных сообщества из которого пришёл запрос;
- обработка запроса – выделение из json структуры необходимых данных, занесение их в отдельные блоки памяти для дальнейшей передачи по системе;
- анализ начислений баллов за задания – система проверяет наличие записей о выполненных заданиях посредством получения значений баллов по выбранному пользователю среди всех транзакций, которые содержатся в базе данных;
- формирование статистики выполненных заданий – система формирует ответ на проверку запрошенных заданий в формате, который указывает клиент;
- передача данных по API ВКонтакте участнику – система создает POST запрос, адресованный на социальную сеть ВКонтакте, в котором содержится информация о статистике и данные получателя сообщения.

Результаты применения подобных систем наблюдаются в социальной сети ВКонтакте в сообществах сервисов ActiveUsers, Автопилот, Senler и др.

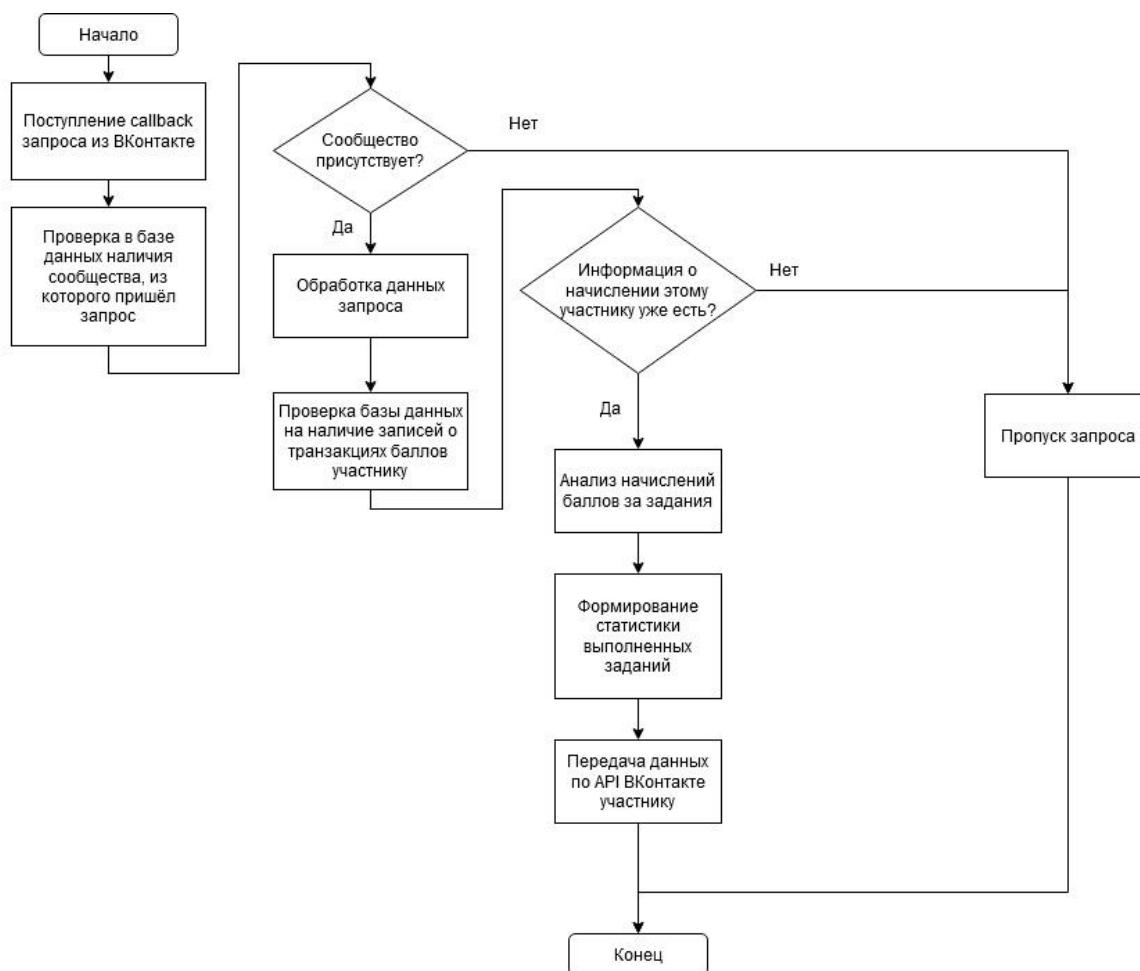


Рис. 4. Блок-схема алгоритма анализа выполнения заданий пользователем

Разработанная информационная система может быть использована владельцами интернет-магазинов или лицами, предоставляющими услуги по сети интернет для внедрения в свои сообщества социальной сети (например, «ВКонтакте») возможности быстрой оплаты заказов участниками сообществ, создавать различные виды игровых механик и получать информацию о статистике различных этапов их реализации, что позволит увеличить количественные показатели продаж.

Библиографический список

1. Friedrich J., Becker M., Kramer F., Wirth M., Schneider M. 2019. Incentive design and gamification for knowledge management. *Journal of Business Research* (in press).
2. Hassan L., Dias A., Hamari J. 2019. How motivational feedback increases user's benefits and continued use: A study on gamification, quantified-self and social networking. *International Journal of Information Management* 46: 151–162.
3. Hofacker C. F., De Ruyter K., Lurie N. H., Manchanda P., Donaldson J. 2016. Gamification and mobile marketing effectiveness. *Journal of Interactive Marketing* 34: 25–36.
4. Хантер Д., К. Вербах. *Вовлекай и властвуй. Игровое мышление на службе бизнеса*; -М., 2016. -224с.
5. Кент Бек *Экстремальное программирование. Библиотека программиста.* – Санкт-Петербург: «Питер», 2015. – 224 с. Г. Зикерман, Д. Линдер. *Геймификация в бизнесе. Как пробиться сквозь шум и завладеть вниманием сотрудников и клиентов*; -М., 2015. -272с.
6. Мэт Зандстра РНР. *Объекты, шаблоны и методики программирования.* – М.: «Вильямс». – 2017. – 560 с.

I.V. Sokolov, Y.S. Yegorov, A.A. Alpeeva

**DEVELOPMENT OF A SALES BUSINESS PROCESS MANAGEMENT SYSTEM
USING GAME METHODS OF USER INVOLVEMENT**

Nizhny Novgorod State Technical University n.a. R.E. Alekseev

Purpose: Develop a business process management system in which elements from the gaming industry can be applied.

Design / methodology / approach: The work showed the process of developing a system from the stage of analyzing the subject area, as well as competitors, to the direct selection of the possibilities for implementing this system. The paragraphs presented in the article are based on time-tested patterns that can be used to further scale the system.

Findings: The article presents the process of analyzing and developing a business process management system using game methods of user engagement.

Research implications/implications: The article examines the current state of business process management systems, their features and disadvantages, and makes a conclusion about the advisability of using such systems.

Originality/value: The materials of the review carried out can be used as a reference for research and develop related to same systems.

Keywords: gamification, business process, information system, marketing, social network.

АРХИТЕКТУРА МНОГОУРОВНЕВОЙ СЕТЕВОЙ СИСТЕМЫ УПРАВЛЕНИЯ УМНЫМИ ФЕРМАМИ

Нижегородский государственный технический университет им. Р.Е. Алексеева

В настоящее время сельское хозяйство определяет сложную управленческую задачу, которая предъявляет строгие требования к информационным системам управления фермерским хозяйством. Предложена архитектура системы управления умной фермой. Промышленный интернет вещей позволяет собирать данные с множества ферм. Внедрение предписывающей аналитики существенно повысит эффективность животноводства.

Ключевые слова: промышленный интернет вещей, мониторинг, большие данные, аналитика, принятие решений, управление умной фермой.

В последнее десятилетие информационные системы в сочетании с облачными вычислениями, мобильными устройствами, сенсорными сетями и Интернетом вещей (IoT) привели к появлению так называемых больших данных, которые характеризуются большими объемом, разнообразием и скоростью поступления (3V). Использование больших данных помогает формировать достаточно точные прогнозы и принимать обоснованные решения [1].

Одна из областей, в которой весьма эффективным представляется внедрение цифровизации, сервисов ИИОТ и интеллектуальных информационно-управляющих систем для обеспечения качественно нового уровня автоматизации представляет сельское хозяйство. Например, птицеводческое хозяйство, как вид деятельности является наиболее требовательным к соблюдению технологических параметров содержания и корма птицы на протяжении всего технологического цикла, начиная от инкубационного периода и до получения товарной продукции (яйцо, мясо). При этом организационно-хозяйственная деятельность птицеводческих ферм содержит большой спектр таких задач, как обеспечение здоровья птиц, управление питанием и микроклиматом, включая автоматическое поддержание параметров температуры и влажности, контроля загрязнения воздуха, а также обеспечение безопасности, включая видеоконтроль и поддержание контрольно-пропускного режима, проведение охранно-пожарных мероприятий и др. За счет цифровизации и поддержания квазиоптимальных режимов рентабельность «умных ферм» может быть повышена примерно на 40% и даже более.

К наиболее востребованным цифровым инструментальным средствам (digital farming) относятся системы управления фермерскими хозяйствами (Farm Management Systems), которые автоматизируют различные производственные процессы, включая управление ресурсами и системы поддержки принятия решений (Decision Support Systems), которые позволяют формировать управленческие решения на основе глубокого анализа «исторических» данных мониторинга.

Крупные хозяйства приступают к внедрению комплексных платформ управления процессами на базе интернета вещей (IoT) – интеллектуальных сетей датчиков, видеокамер, беспилотных аппаратов и других устройств, робототехники, облачных технологий искусственного интеллекта (ИИ). В частности, платформа Episcopy [2], разработанная компанией WaveAccess на базе технологий IoT и ИИ, помогает анализировать параметры, влияющие на эффективность ферм и прогнозировать производственные показатели. Датчики, установленные в помещениях фермы, фиксируют параметры процессов, такие как данные по потреблению корма и микроклимату, включая температуру, влажность, давление, с использованием которых рассчитываются показатели эффективности функционирования.

Повышение эффективности ферм и птицефабрик возможно за счет внедрения современных систем мониторинга и автоматического управления, внедрение которых позволяет построить так называемую «умную ферму» [3]. Например, от качественно приготовленного комбикорма зависит более 60% рентабельности всего производства [4]. Использование дозаторов высокого класса точности повышает соответствие состава рецептуры. При этом появляется возможность выявления зависимостей показателей качества функционирования ферм и птицефабрик от особенностей рациона кормления и других параметров. В частности, на прирост живой массы и повышение яйценоскости кур непосредственное влияние оказывают параметры среды обитания, такие как состав воздуха, температура, влажность и т.д. Значения этих параметров, состава кормов и др., используемые в различных хозяйствах, могут отличаться. Объединение в сеть различных ферм или птицефабрик, сбор данных об используемых параметрах и достигаемых показателях позволит с помощью процедур дескриптивной аналитики синтезировать модели влияния. Используя эти модели, предписывающая аналитика обеспечит формирование рекомендаций по управлению режимами функционирования для множества ферм и птицефабрик, объединенных сетью промышленного интернета вещей (IIoT).

В настоящее время разнообразные технологические объекты управления (ТОУ), в том числе в сельскохозяйственной сфере, управляются с помощью собственных (автономных) автоматизированных систем управления технологическими процессами (АСУ ТП), которые обычно включают три иерархических уровня. На верхнем уровне решаются задачи диспетчеризации процессов, как правило, ручного управления режимами, расчета основных технико-экономических показателей производства, визуализации данных, ведения журналов и формирования отчетов, а также диагностики. На среднем уровне, реализуемом на базе программируемых логических контроллеров (ПЛК), решаются задачи автоматического управления и регулирования, такие как пуск и останов оборудования, логико-командное управление, аварийное отключение и защита. Нижний (полевой) уровень обеспечивает сбор данных о параметрах технологического процесса и состояния оборудования с помощью разнообразных датчиков и реализует управляющие воздействия с помощью исполнительных устройств.

Например, в составе АСУ ТП современной птицефермы можно выделить следующие основные подсистемы [3]:

- подсистема управления микроклиматом обеспечивает автоматическое поддержание параметров температуры, влажности и чистоты (контроля загрязнения) воздуха;
- подсистема управления хранением и подачей корма отслеживает заполнение внешнего бункера, контролирует хранение корма и управляет загрузочным люком;
- подсистема управления поением выполняет автоматическое манипулирование лебедками поения;
- подсистема управления внутренним и внешним освещением обеспечивает автоматическое включение, выключение и поддержание требуемого уровня освещенности птичника;
- подсистема видеоаналитики позволяет контролировать возникновение нештатных ситуаций, способствуя значительному снижению риска аварий.

Внедрение дополнительных систем видеоаналитики и бесконтактного дистанционного контроля состояния и здоровья животных позволит повысить информативность данных мониторинга. Непрерывный мониторинг и накопление информации в базе данных позволит наряду с оперативным контролем отклонений от заданных параметров проводить аналитическую обработку данных и оптимизировать функционирование за счет управления режимами. С другой стороны, система мониторинга позволяет контролировать и управлять техническим состоянием оборудования, обеспечивая повышение показателей безотказности.

Применение технологий промышленного интернета вещей позволяет перейти от независимого управления отдельными ТОУ к управлению совокупностью ТОУ. Обработка статистического ансамбля реализаций, формируемых от совокупности типовых объектов управления позволяет значительно снизить уровень неопределенности и получить потенциальный выигрыш по точности управления.

Таким образом, предлагается использовать технологии промышленного интернета вещей для объединения множества территориально распределенных ТОУ в одну систему, что позволит централизованно обрабатывать данные мониторинга и кардинально повысить качество управления (рис. 1а). Вместе с тем, реализация сетевого управления сопряжена со следующими факторами: ограниченная пропускная способность каналов связи; задержки и возможные потери информационных пакетов; несократимая вероятность искажения управляющих воздействий, в том числе в связи с угрозами кибербезопасности.

В связи с этим предлагается компромиссный подход, позволяющий преодолеть ряд указанных ограничений и сдерживающих факторов на основе централизации части функций верхнего уровня АСУ ТП для достижения экстремальных значений принятого критерия качества посредством поддержки диспетчерских решений на начальном этапе развертывания системы. Затем по мере накопления знаний и валидации реализуемых алгоритмов может быть выполнен переход от режима рекомендательной системы к высокоавтоматизированному применению команд управления режимами на уровне локальных АСУ ТП.

Для разработки решающих правил в квазистатическом приближении может быть сформулирована задача оптимизации, решение которой представляют величины уставок, обеспечивающие максимизацию обобщенных показателей функционирования ТОУ (ферм или птицефабрик). Однако наличие неопределенности стохастического характера, вызванной действием ряда возмущений и неконтролируемых факторов, отсутствие точных моделей ТОУ, проявление нелинейностей, а также невысокая согласованность экспертных суждений значительно повышает сложность решаемой задачи. Всё сказанное приводит к целесообразности применения на верхнем уровне предложенной архитектуры (рис. 1а) интеллектуальной информационно-управляющей системы (ИИУС), которая может быть построена с применением технологий интеллектуального анализа данных на основе решений, предложенных в [5-6].

Для обработки накапливаемых «исторических» данных о функционировании системы и синтеза моделей ТОУ наиболее проработанным подходом является применение параметрических и непараметрических методов идентификации, а также методов структурной идентификации. Однако использование традиционных методов может быть проблематичным, поскольку реализация идентификационного синтеза характеризуется заметными трудностями. Прежде всего, при совмещении управления и оценивания управляемый процесс оказывается нестационарным с зависимыми значениями, поэтому требование состоятельности оценок ограничивает класс стратегий управления. Помимо этих трудностей для широкого класса объектов построение классических идентификационных моделей оказывается практически невозможным. Альтернативой могут служить методы адаптации, использующие индуктивные знания, получаемые в результате анализа данных [7].

Представление данных мониторинга (рис. 1б), поступающих в ИИУС, основано на реализации концепции многомерности и иерархичности [8]. Для визуализации данных мониторинга в сложных распределенных системах могут быть использованы такие измерения, как «пространство» и «время».

Измерение «пространство» понимается в обобщенном смысле и включает иерархические уровни, соответствующие отдельным ТОУ, их подсистемам и элементам.

Измерение «время» (ретроспектива, текущий момент, прогноз) включает иерархические уровни представления состояния ТОУ за сутки с детализацией по часам и минутам; за месяц с детализацией по дням; за несколько лет с детализацией по месяцам.

Качество функционирования распределенных ТОУ характеризуется, как правило, несколькими показателями $J_i, i = \overline{1, q}$. Если при управлении режимами ТОУ эти показатели используются в качестве частных целевых функций, то задача приобретает многокритериальный характер. Наиболее простой способ сведения многокритериальной задачи к однокритериальной представляет свертка частных целевых функций, называемая также сверткой критериев. При этом обоснование выбора вида свертки, среди которых наиболее популярными являются аддитивная и мультипликативная свертка, представляет собой отдельную задачу.

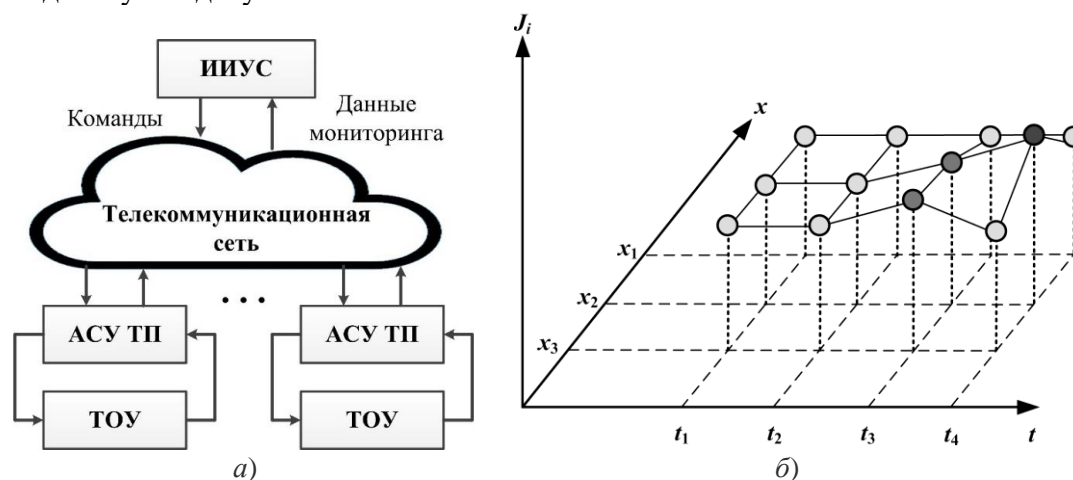


Рис. 1. Архитектура многоуровневой сетевой системы управления (а) и структура данных мониторинга (б)

Ожидается, что предложенная архитектура, основанная на согласованной обработке данных мониторинга от множества ТОУ и на введении дополнительного уровня интеллектуального управления, позволит реализовать концепцию умной фермы и может быть востребована при проведении модернизации средних и крупных фермерских хозяйств и особенно агрохолдингов.

Библиографический список

1. Шамасна Х.А., Андре Н.Г. Роль аналитики больших данных в промышленном интернете вещей // Информационные системы и технологии (ИСТ-2020): Сборник материалов XXVI Междунар. науч.-техн. конф. – Н. Новгород: НГТУ им. Р.Е. Алексеева, 2020, С. 228-232.
2. Аналитическая платформа Episcopo [Электронный ресурс] URL:<https://globalcio.ru/live/projects/3370/>
3. Третьяков С.А., Шержуков Е.Л. Умная птицеферма: взгляд в ближайшее будущее// Автоматизация в промышленности. 2009. № 1. С. 41 – 44.
4. Диспетcharизация в сельском хозяйстве [Электронный ресурс] URL:<https://smekni.com/a/121476/dispetcharizatsiya-v-selskom-khozyaystve/>
5. Милов В.Р., Сулов Б.А., Крюков О.В. Интеллектуализация поддержки управленческих решений в газовой отрасли// Автоматизация в промышленности. 2009. № 12. С. 16-20.
6. Баранов В.Г., Милов В.Р., Сулов Б.А. Способы аналитической обработки информации для системы поддержки принятия решений// Информационно-измерительные и управляющие системы. 2010. Т. 8. № 2. С. 12-15.
7. Теория управления (дополнительные главы): Учебное пособие / Под ред. Д. А. Новикова. –М.: ЛЕНАНД, 2019. –552 с.
8. Милов В.Р., Цыс В.М., Егоров Ю.С., Алексеев В.В., Севрюков А.А. Способ организации доступа и визуализации данных мониторинга распределенных технических объектов // Информационно-измерительные и управляющие системы. № 7. 2013. С. 59-66.

H.A. Shamasna

**AN ARCHITECTURE OF SMART FARM MULTILAYER NETWORKED
MANAGEMENT SYSTEM**

Nizhny Novgorod State Technical University n.a. R.E. Alekseev

Farming nowadays is a complex managerial task that imposes stringent requirements on farm management information systems. The architecture of the smart farm management system is proposed. The industrial internet of things allows to collect data from many farms. Applying of prescriptive analytics will significant improve the efficiency of animal husbandry.

Keywords: industrial internet of things, monitoring, big data, analytics, decision making, smart farm management

АНАЛИЗ РАЗВИТИЯ СИСТЕМЫ ВИДЕОНАБЛЮДЕНИЯ ПРИ ПОСТРОЕНИИ АППАРАТНО-ПРОГРАММНОГО КОМПЛЕКСА «БЕЗОПАСНЫЙ ГОРОД» В НИЖЕГОРОДСКОЙ ОБЛАСТИ

Нижегородский государственный технический университет им. Р. Е. Алексеева

Одной из актуальных задач, возникающих при построении аппаратно-программного комплекса «Безопасный город» в Нижегородской области, является необходимость агрегирования видеопотоков с различных систем видеонаблюдения. В ходе изучения возможных вариантов агрегирования видеопотоков авторами выбран оптимальный с точки зрения системного анализа.

Ключевые слова: агрегирование видеопотоков, интегрирование систем видеонаблюдения.

Введение

Используемые сокращения:

СВН - система видеонаблюдения;

ССВН - сторонняя система видеонаблюдения;

СВН ЦБДД - система видеонаблюдения ГКУ НО «ЦБДД».

С целью повышения эффективности обеспечения комплексной безопасности населения Нижегородской области при проведении Чемпионата Мира по футболу FIFA-2018 (ЧМ-2018), Правительство Нижегородской области заключило соглашение №65-П от 07 апреля 2017 года «О создании комплекса средств автоматизации аппаратно-программного комплекса «Безопасный город» Нижегородской области и развитии существующей системы видеонаблюдения (СВН) правоохранительного сегмента и существующей системы фото-видео фиксации нарушений правил дорожного движения в Нижегородской области» [4].

Аппаратно-программный комплекс «Безопасный город» включает в себя системы автоматизации деятельности единой дежурно-диспетчерской службы, муниципальных служб различных направлений, системы приема и обработки сообщений, системы обеспечения вызова экстренных и других муниципальных служб различных направлений деятельности, системы мониторинга, прогнозирования, оповещения и управления всеми видами рисков и угроз, свойственных данному муниципальному образованию [2].

Видеонаблюдение вносит весомый вклад в обеспечение безопасности населения, поэтому предусмотрена ее интеграция в аппаратно-программный комплекс «Безопасный город» [3].

До 2012 года строительство систем видеонаблюдения в Нижегородской области осуществлялось фрагментировано. В результате в Нижегородской области развернуты системы видеонаблюдения, которые не придерживаются общих или открытых стандартов, что затрудняет их интеграцию.

С целью выполнения государственных функций и иных полномочий в сфере обеспечения безопасности дорожного движения и общественной безопасности Правительством Нижегородской области в ноябре 2012 года было принято решение о создании государственного казенного учреждения Нижегородской области «Центр безопасности дорожного движения» (далее ГКУ НО «ЦБДД»). В рамках реализации указанного решения на территории Нижегородской области установлено: 225 стационарных комплексов фотовидеофиксации нарушений ПДД (фиксация скоростного режима – 195, фиксация правил проезда перекрестков – 20 комплексов, на постах ДПС – 5 комплексов, фиксация нарушений правил проезда на железнодорожных переездах – 5 комплексов); 602 камеры видеонаблюдения, расположенные во всех районах города Нижнего Новгорода [1].

Варианты агрегирования видео потоков

С целью реализации единого системного подхода к обеспечению общественной безопасности, правопорядка и безопасности среды обитания необходимо интегрировать СВН различных объектов региона (спортивные объекты, объекты транспортной инфраструктуры, гостиницы, торговые центры и т.д.) [3].

Однако, с учетом того, что на каждом объекте внедрено программное обеспечение определённого производителя, интеграция локальной СВН объекта в СВН АПК «Безопасный город» в большинстве случаев не представляется возможной. С целью создания единой системы видеонаблюдения в рамках развития программы «Безопасный город» на территории Нижегородской области проведен анализ получения видеопотоков в систему (ГКУ НО «ЦБДД») из сторонних систем видеонаблюдения по IP-сетям.

С целью получения видеопотоков в систему видеонаблюдения ГКУ НО «ЦБДД» с видеокамер ССВН используются общепринятые протоколы обмена: ONVIF, HTTP, RTSP.

Далее приведены варианты интеграции видеопотоков в систему видеонаблюдения ГКУ НО «ЦБДД» из сторонних систем видеонаблюдения по IP-сетям.

1. Получение видеопотоков с камер ССВН при непосредственном подключении к камерам ССВН изображено на рис. 1.

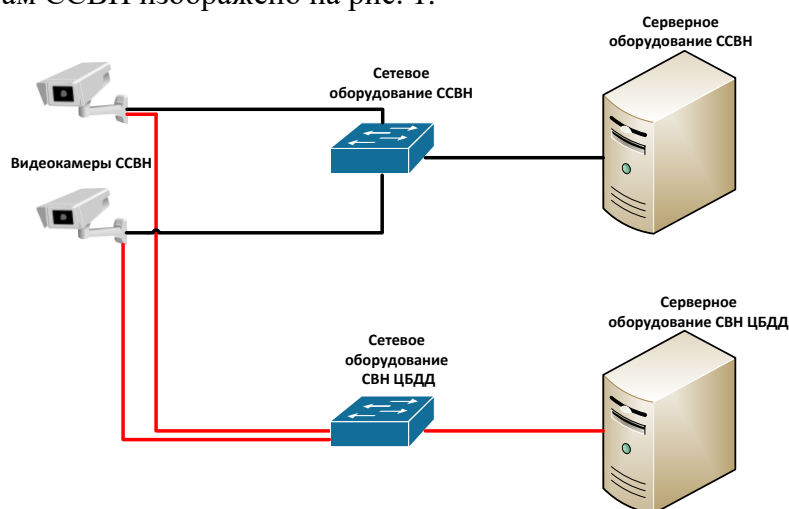


Рис. 1. Непосредственное подключение к видеокамерам

Каждая камера формирует два видеопотока высокого качества. Один видеопоток через сетевое оборудование ССВН доставляется к серверам ССВН, другой видеопоток через сетевое оборудование СВН ЦБДД доставляется к серверам СВН ЦБДД.

Достоинства:

- отсутствует необходимость в приобретении со стороны ГКУ НО «ЦБДД» дополнительных лицензий на программное обеспечение ССВН.

Недостатки:

- ГКУ НО «ЦБДД» вынуждено приспособливаться к IP-адресному пространству ССВН;
- отсутствует доступ к видеоархиву ССВН, что приводит к необходимости увеличения вычислительной мощности видеосерверов и емкости систем хранения данных со стороны ГКУ НО «ЦБДД»;
- в камерах видеонаблюдения отсутствует возможность назначения приоритетов управления для операторов из различных СВН, что снижает эффективность их использования;
- не все модели камер имеют по два сетевых интерфейса и вычислительную мощность, позволяющую формировать 2 видеопотока высокого качества;

- достаточно высокие финансовые затраты на организацию участка «последней мили»;
 - сетевой трафик с камер ССВН увеличивается в 2 раза. В результате может возникать перегрузка сети ССВН, в следствие чего собственник ССВН будет вынужден увеличивать скорость каналов связи (в т.ч. за счет увеличения финансирования арендованных каналов) и вычислительные мощности сетевого оборудования.
2. Получение видеопотоков с камер ССВН при подключении к сетевой части ССВН изображено на рис. 2.

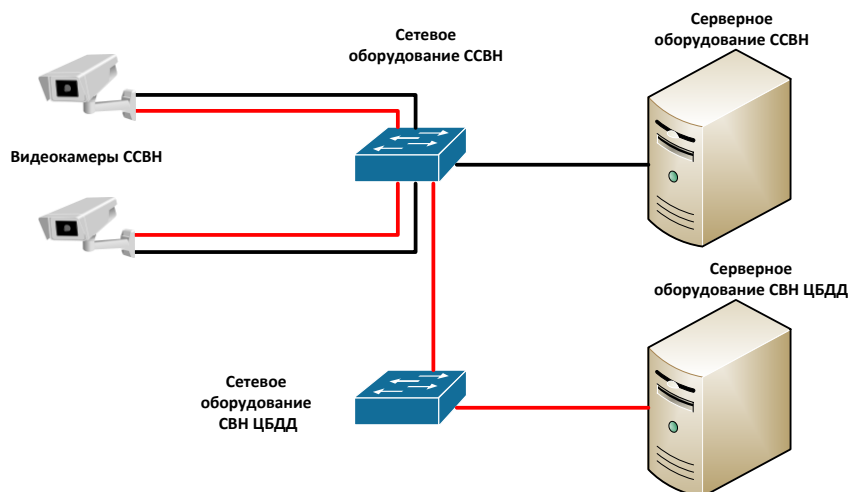


Рис. 2. Подключение к сетевой части ССВН

Каждая камера формирует два видеопотока высокого качества (один видеопоток - к серверу ССВН, другой - к серверу СВН ЦБДД). Все видеопотоки концентрируются на агрегирующем коммутаторе ССВН и далее разделяются по соответствующим направлениям. Требуется подключение на агрегирующем коммутаторе ССВН, который объединяет все видеопотоки, что снижает финансовые затраты на организацию «последней мили».

Достоинства:

- отсутствует необходимость в приобретении со стороны ГКУ НО «ЦБДД» дополнительных лицензий на программное обеспечение ССВН;
- по сравнению с предыдущим способом отсутствуют затраты на организацию «последней мили».

Недостатки:

- ГКУ НО «ЦБДД» вынуждено приспосабливаться к IP-адресному пространству ССВН;
- камеры видеонаблюдения, как и в первом варианте, передают 2 видеопотока, что увеличивает в 2 раза сетевой трафик. В результате может возникать перегрузка сети ССВН, в следствие чего собственник ССВН будет вынужден увеличивать скорость канала связи и вычислительные мощности сетевого оборудования;
- не все модели камер имеют вычислительную мощность, позволяющую формировать 2 видеопотока высокого качества;
- отсутствует доступ к видеоархиву ССВН, что приводит к необходимости увеличения вычислительной мощности видеосерверов и емкости систем хранения данных со стороны ГКУ НО «ЦБДД»;
- в камерах видеонаблюдения отсутствует возможность назначения приоритетов управления для операторов из различных СВН, что снижает эффективность их использования.

3. Получение видеопотоков с камер ССВН при подключении к видеосерверу ССВН изображено на рис. 3.

Каждая камера формирует один видеопоток высокого качества и передает его к серверам ССВН. Подключение серверов СВН ЦБДД осуществляется непосредственно к видеосерверу ССВН через сетевое оборудование.

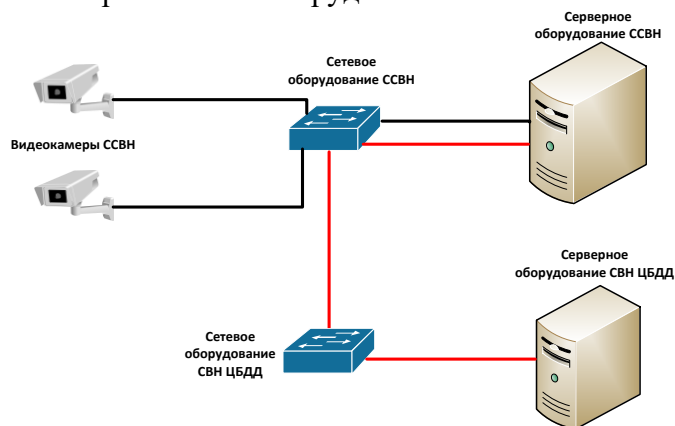


Рис. 3. Подключение к видеосерверу ССВН

Достоинства:

- минимизация затрат на организацию «последней мили»;
- есть возможность управления камерами ССВН и назначения приоритетов управления для операторов из различных СВН;
- есть доступ к видеоархиву ССВН;
- отсутствует нагрузка на камеры и затраты на расширение каналов связи до камер ССВН.

Недостатки:

- в ССВН используется программное обеспечение отличное от программного обеспечения, используемого в СВН ЦБДД. Для их интеграции необходимо создание и использование единого стека протоколов/интерфейсов;
- возможно потребуется расширение лицензионных ограничений на серверном оборудовании ССВН.
- Получение видеопотоков с камер ССВН при подключении к видеосерверу ССВН с помощью автоматизированного рабочего места ГКУ НО «ЦБДД» изображено на рис. 4.

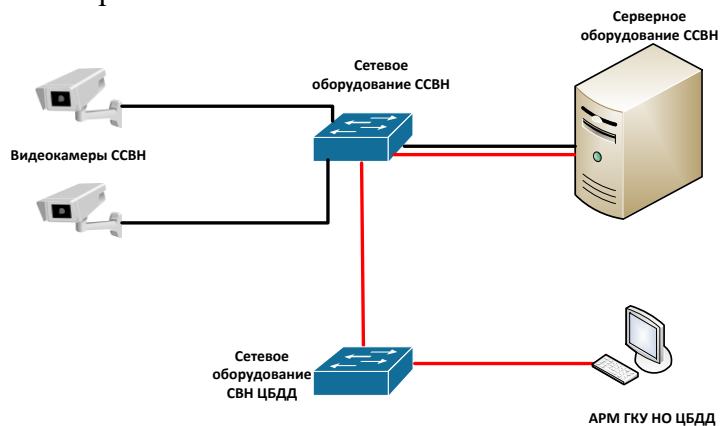


Рис. 4. Подключение к видеосерверу ССВН с помощью автоматизированного рабочего места ГКУ НО «ЦБДД»

Каждая камера формирует один видеопоток высокого качества и передает его к серверам ССВН. осуществляется непосредственное подключение арм гку но «цбдд» к видеосерверам ССВН через сетевое оборудование.

Достоинства:

- минимальные затраты на организацию подключения к ССВН;
- отсутствуют затраты на организацию «последней мили»;
- есть возможность управления камерами ССВН и назначения приоритетов управления для операторов из различных СВН;
- есть доступ к видеоархиву ССВН;
- отсутствует нагрузка на камеры и затраты на расширение каналов связи до камер ССВН.

Недостатки:

- отсутствует возможность объединения нескольких систем видеонаблюдения на одном рабочем месте;
- не масштабируемое решение.

Заключение

Указанные требования нужно применять при формировании технических заданий на создание, развитие и эксплуатацию информационных систем и программно-технических комплексов видеонаблюдения, входящих в состав АПК «Безопасный город», либо интегрируемых с ним.

Реализация системного подхода к обеспечению общественной безопасности, правопорядка и безопасности среды обитания позволит полностью реализовать потенциал видеонаблюдения, а именно быстрее и проще осуществлять доступ к видеопотоку, управлять камерами видеонаблюдения и получать доступ к видеоархиву ССВН.

Библиографический список

1. Государственное казенное учреждение Нижегородской области "Центр безопасности дорожного движения" [Электронный ресурс]. – URL: <http://gkunobg.ru/about/osnovnye-tseli>(дата обращения: 26.05.2018).
2. Единые требования к техническим параметрам сегментов аппаратно-программного комплекса "Безопасный город" [Электронный ресурс]. – URL: http://07.mchs.gov.ru/upload/site38/document_file/N0LDsp5FVR.pdf(дата обращения: 26.05.2018).
3. Об утверждении Концепции построения и развития аппаратно-программного комплекса «Безопасный город»: Распоряжение Правительства Российской Федерации от 3 декабря 2014 года № 2446-р [Электронный ресурс]. – URL: <http://government.ru/media/files/OapVppc8jyA.pdf> (дата обращения: 26.05.2018).
4. Соглашение о создании комплекса средств автоматизации аппаратно-программного комплекса «Безопасный город» Нижегородской области и развитии существующей системы видеонаблюдения правоохранительного сегмента и существующей системы фото-видеофиксации нарушений правил дорожного движения в Нижегородской области № 65-П от 07.04.2017, приложение 1 – 20, 21 с.

O. E. Vinogradov, A. V. Semashko

**ANALYSIS OF THE DEVELOPMENT OF THE VIDEO SURVEILLANCE SYSTEM IN
THE CONSTRUCTION OF THE HARDWARE AND SOFTWARE COMPLEX "SAFE
CITY" IN THE NIZHNY NOVGOROD REGION**

Nizhny Novgorod State Technical University n. a. R. E. Alekseev

One of the urgent tasks that arise when building the hardware and software complex "Safe City" in the Nizhny Novgorod region is the need to aggregate video streams from various video surveillance systems. In the course of studying possible options for aggregating video streams, the authors selected the optimal one from the point of view of system analysis.

Keywords: aggregation of video streams, integration of video surveillance systems.

УСТРАНЕНИЕ ИНВЕРСИИ В ЦИКЛИЧЕСКИХ КОДАХ ДЛЯ ПОВЫШЕНИЯ ПОМЕХОУСТОЙЧИВОСТИ СИСТЕМ СВЯЗИ

Нижегородский государственный технический университет им. Р.Е. Алексеева

Рассмотрен метод двоичного помехоустойчивого кодирования с возможностью обнаружения инверсии всех символов в кодовом блоке. Путем имитационного моделирования проведено исследование кодов на базе циклических, позволяющих кроме одиночных ошибок обнаруживать и исправлять инверсию символов кодового блока. Исследовались коды Хемминга (15, 11) и код БЧХ (31, 16) в программной среде Matlab.

Ключевые слова: моделирование, циклические коды, помехоустойчивое кодирование, системы связи, система управления, инверсия кодового блока.

В настоящее время по каналам связи передаются цифровые данные со столь высокими требованиями к достоверности передаваемой информации, что удовлетворить эти требования путём совершенствования антенно-фидерных трактов радиолиний, увеличения излучаемой мощности, снижения собственного шума приемника оказывается экономически невыгодным или просто невозможным.

Несмотря на то, что современные системы передачи данных отвечают всем основным стандартам и требованиям, они все еще далеки от совершенства. Это связано с очевидными требованиями увеличения скорости передачи сообщений по каналам коммуникации, и, в связи с этим, возникновением искажений сигналов, которые не встречаются при меньших скоростях. Улучшение качества связи является важнейшей задачей настоящего времени. Для повышения помехоустойчивости используют множество различных способов. Применение помехоустойчивых кодов, лежащих в основе устройств кодирования и декодирования, является одним из средств решения данной задачи.

Основные функции цифровой системы связи обычно разделяют, как показано на блок-схемах на рис. 1 [1]. Данные поступающие в систему связи от источника информации обрабатываются кодером источника, предназначенным для более компактного представления данных источника. Затем данные обрабатываются кодером канала, преобразующим последовательность символов кодового слова источника в другую последовательность символов, называемую кодовым словом канала. Кодовое слово канала – это новая, более длинная последовательность с большей избыточностью, чем в исходном коде. При этом введенная кодером канала избыточность является регулярной. Каждый символ кодового слова канала может быть представлен битом или группой бит, в зависимости от применяемой в кодере алгебры.

Модулятор преобразует каждый символ кодового слова канала в соответствующий аналоговый сигнал из конечного множества допустимых сигналов. По каналу передается последовательность этих сигналов. Так как в канале возникают различного типа шумы и искажения, то процесс на выходе канала отличается от входного. Демодулятор преобразует каждый сигнал, полученный на выходе канала, в последовательность символов одного из кодовых слов канала (формируется оценка). При этом на выходе демодулятора могут возникать ошибки.

Декодер источника выполняет операцию, обратную операции кодера источника, результат которой поступает к получателю [1,3].

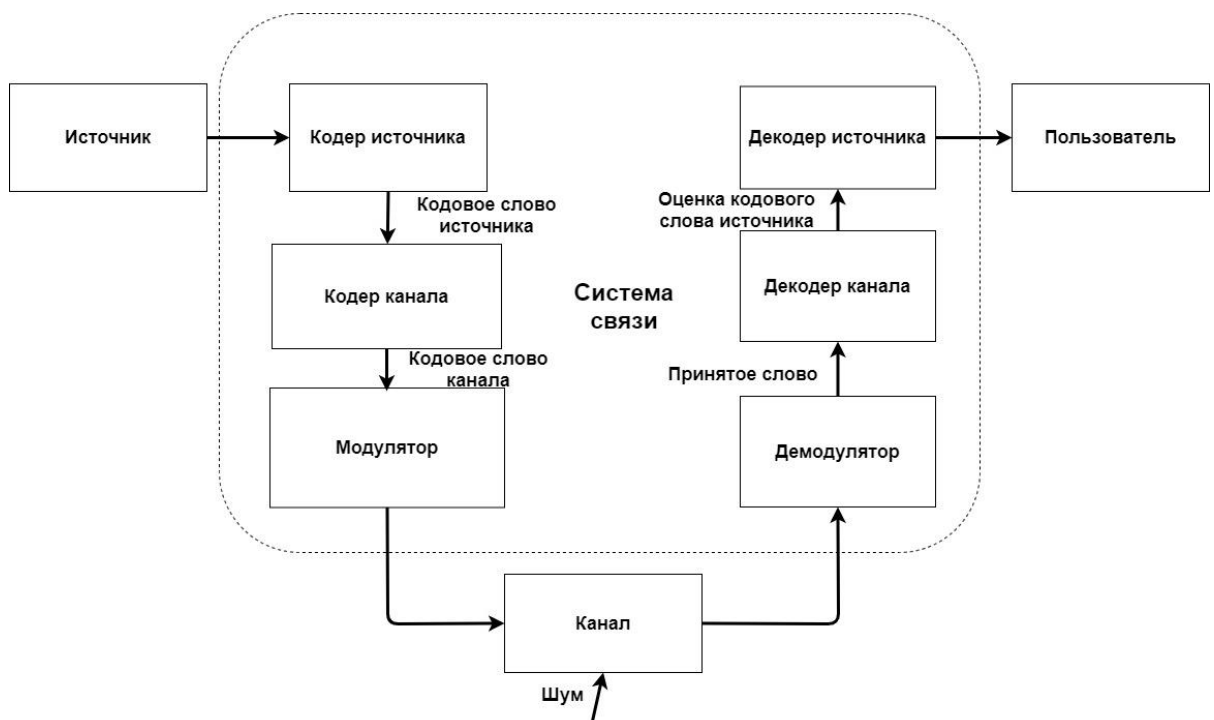


Рис. 1. Модель канала

Вместе с тем в двоичном канале связи может происходить инверсия двоичных символов кодового слова. Это явление, например, может возникать при двухуровневой фазовой модуляции, когда из-за «перескока» фазы может возникать эффект «обратной работы», когда ошибка в одном символе может привести к ошибочному детектированию всех последующих символов, это и приводит к инверсии [2].

Эта проблема имеет различные решения. Ниже приводится метод борьбы с эффектом «обратной работы» на основе корректирующих кодов. Для проверки правильности предложенных решений приводятся результаты моделирования в среде Matlab кодов на базе циклических, в частности кодов Хемминга и кодов Боуза-Чоудхури-Хоквингема (БЧХ).

Коды Хэмминга – простейшие линейные коды с минимальным расстоянием 3, способные исправить одиночную ошибку. Циклическое представление кодов Хемминга является частным случаем кодов БЧХ [3]. Одно из преимуществ циклических кодов состоит в том, что существует несколько эффективных алгоритмов их декодирования (Берлекэмп-Мэсси, Питерсона-Горенштейна-Цирлера, на основании алгоритма Евклида и так далее) [6].

Данные коды имеют возможность вычисления контрольной суммы (защитных символов) – значение, рассчитываемое по определенной схеме на основе закодированного сообщения. Если исходные и рассчитанные суммы не равны, принимается решение о ненадежности полученных данных (обнаружение ошибок), и пакет может быть запрошен для повторной передачи. Пример алгоритма расчета контрольной суммы приведен на рис. 2. Сначала сообщение дополняется нулями до определенной длины (зависит от используемых кодов). Потом при помощи операции «побитовое или» сообщение делится на образующий полином, и остаток от деления ставится на место нулей. При проверке на стороне получателя сообщение так же делится на образующий многочлен. Если остаток от деления есть ноль, то считается, что сообщение пришло без ошибки. Если остаток от деления есть комбинация нулей и единиц, то сообщение пришло с ошибкой, и формируется синдром ошибки [3,5].

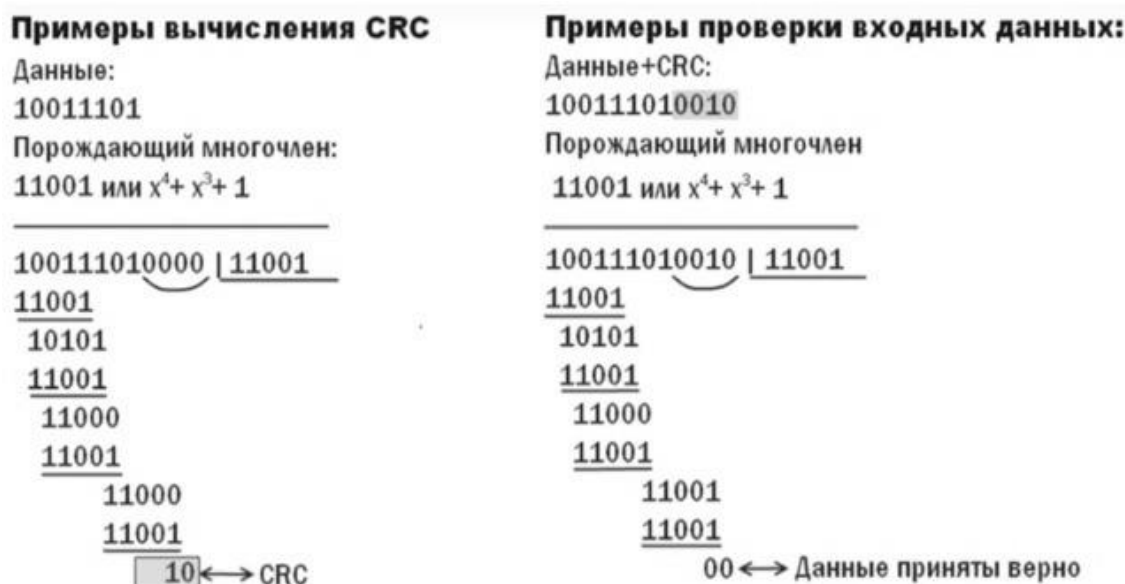


Рис. 2. Алгоритм расчета контрольной суммы

Метод борьбы с инверсией основан на следующем подходе. Поскольку контрольная сумма не позволяет определить инверсию, нужно создать дополнительный блок символов, который будем называть «блок проверки инверсии». Этот блок будет определять инверсию, и тогда, после повторной инверсии контрольная сумма станет доступной для работы.

Определение инверсии кодового блока представляет собой отдельную задачу, поскольку сам кодовый блок может содержать ошибки. Показано, что если в кодовом блоке не более одной ошибки, то достаточно 3 дополнительных битов, позволяющих определять инверсию. Если в кодовом блоке до двух ошибок, то для определения инверсии необходимо 5 бит и так далее.

Сформулируем правила работы блока проверки инверсии длиной 3 бита:

1. Если на выходе демодулятора количество «нулей» равно двум или трем, то инверсия кодового блока не произошла.
2. Если количество «единиц» равно двум или трем, то произошла инверсия.

В качестве примера рассмотрим ситуацию для кодов Хемминга, когда приходит сообщение длиной 15 символов, в которых так же содержится контрольная сумма. На рис. 3 продемонстрировано то, как нужно расположить блок проверки инверсии.

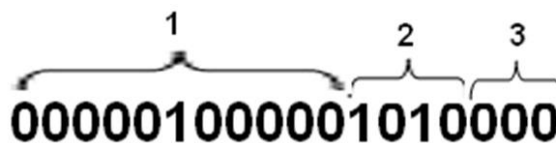


Рис. 3. Схема расположения блока проверки инверсии

Описание алгоритма работы обнаружения ошибок:

1. Проверяется инверсия сообщения.
2. Проверяется наличие ошибки в блоке проверки инверсии.
3. Если инверсия сообщения не произошла, контрольная сумма проверяет целостность информационного слова и самой контрольной суммы.
4. Если произошла инверсия сообщения, проводится обратная инверсия, и осуществляется декодирование в обычном режиме.

Применение блока проверки инверсии, длиной 3 бита позволяет решить проблему обратной работы демодулятора.

В работе рассматривается процедура имитационного моделирования, рассмотренного выше метода в среде MATLAB. MATLAB – язык программирования высокого уровня, позволяющий быстро выполнять вычисления для больших таблиц данных, что значительно упрощает моделирование.

Моделирование радиосистем проводилось в программной среде MATLAB – Simulink. Simulink позволяет использовать готовую библиотеку блоков для моделирования электрических, механических и гидравлических систем, а также применять модельный подход к разработке систем управления, цифровой связи и устройств в режиме реального времени. Simulink также имеет библиотеки для тестирования различных типов кодирования, используемых при моделировании радиосистем. Коды могут быть классифицированы в соответствии с различными критериями. Главная особенность MATLAB – это широкие возможности работы с массивами [6].

На рис. 4 приведен алгоритм пошагового моделирования циклических кодов для повышения помехоустойчивости систем связи.

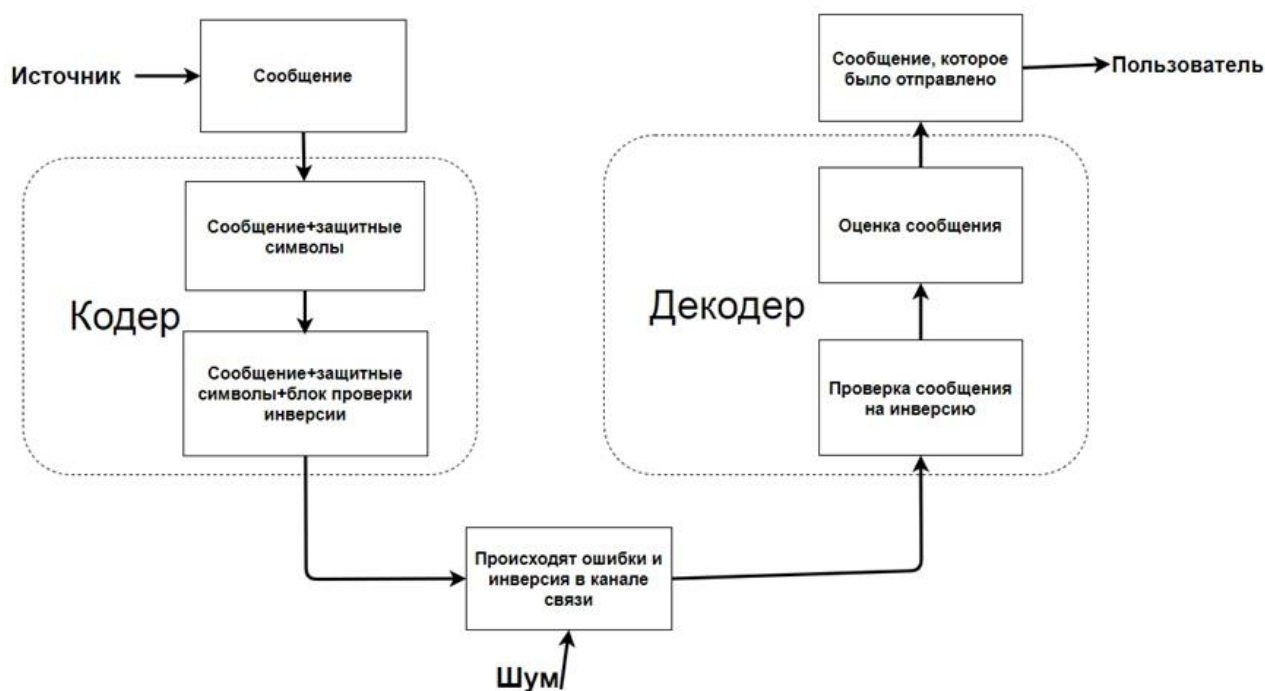


Рис. 4. Алгоритм пошагового моделирования циклических кодов для повышения помехоустойчивости систем связи

На основе полученных алгоритмов в программе MATLAB была проведена модуляция следующих 8 ситуаций, в соответствии со схемой на рис. 5:

- Моделирование сообщения (15,11), ошибочной инверсии и проверки/исправления инверсии;
- Моделирование сообщения (15,11), при отсутствии ошибочной инверсии и проверки/исправления инверсии;
- Моделирование сообщения (15,11), ошибочной инверсии и проверки/исправления инверсии с учетом одной ошибки в блоке проверки инверсии;
- Моделирование сообщения (15,11), при отсутствии ошибочной инверсии и проверки/исправления инверсии с учетом одной ошибки в блоке проверки инверсии;

- Моделирование сообщения (31,16), ошибочной инверсии и проверки/исправления инверсии;
- Моделирование сообщения (31,16), при отсутствии ошибочной инверсии и проверки/исправления инверсии;
- Моделирование сообщения (31,16), ошибочной инверсии и проверки/исправления инверсии с учетом одной ошибки в блоке проверки инверсии;
- Моделирование сообщения (31,16), при отсутствии ошибочной инверсии и проверки/исправления инверсии с учетом одной ошибки в блоке проверки инверсии.

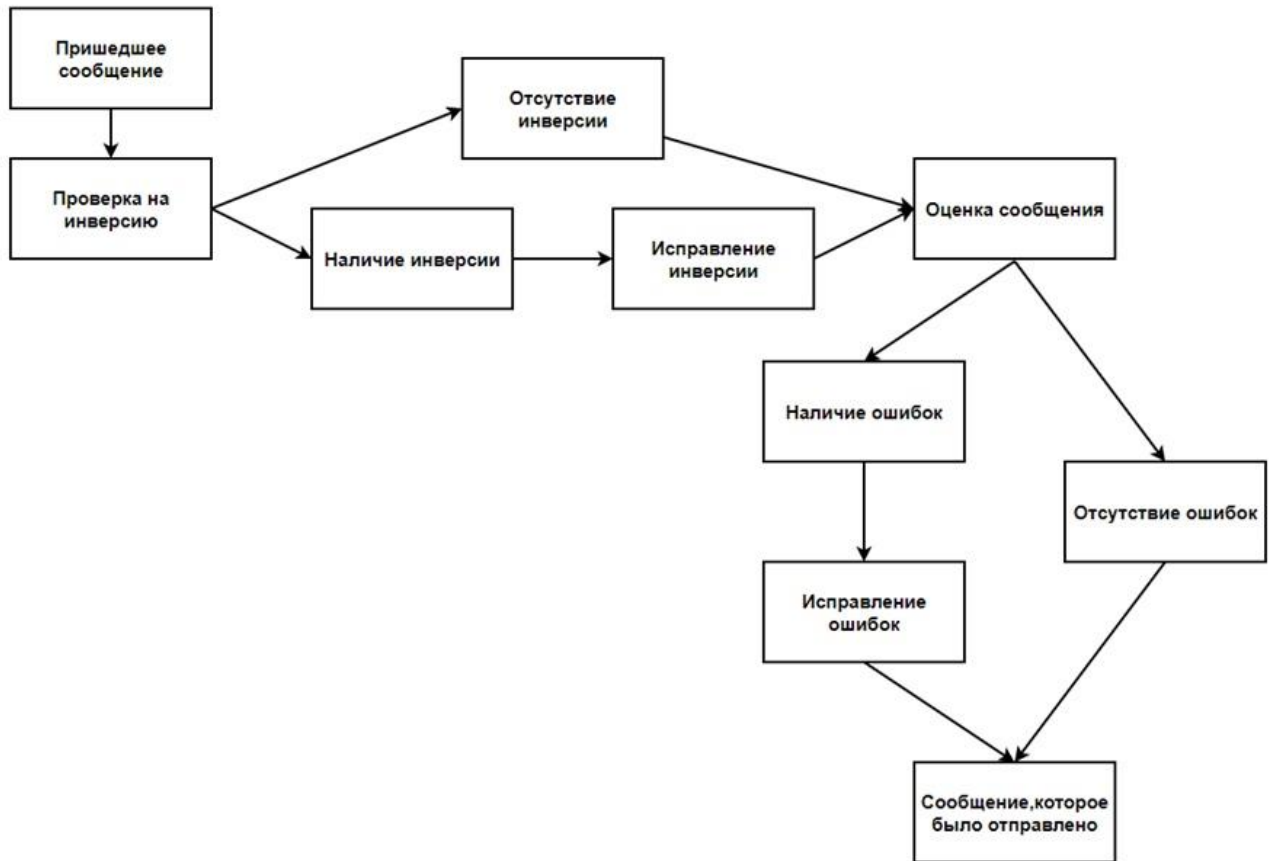


Рис. 5. Алгоритм обнаружения и исправления ошибок в циклических кодах

Моделирование показало, что данный метод является рабочим для распознавания и исправления инверсии циклического кода сообщения в двоичном канале связи для кодов Хемминга и БЧХ-кодов.

Библиографический список

1. Шеннон К.. Работы по теории информации и кибернетики.[Сборник статей] : Пер. с англ. / С предисл. А. Н. Колмогорова ; Под ред. Р. Л. Добрушина и О. Б. Лупанова. - Москва : Изд-во иностранной литературы, 1963. - 829 с.
2. Финк Л.М. Теория передачи дискретных сообщений. Изд.2е, переработанное, дополненное. – Изд-во "Советское радио", – 1970. – 728 с.
3. Блейхут Р. «Теория и практика кодов, контролирующих ошибки. – Theory and Practice of Error Control Codes.» – М.: Мир, 1986. – 576 с.
4. Жельников В. «Криптография от папируса до компьютера». – М.: АБФ, 1996 – 355 с.

5. Коккоз М.М., Ж. Сайлаукызы «Исследование корректирующей способности Кода Хэмминга» – Информационный издательский учебно-научный центр "Стратегия будущего" (Санкт-Петербург). ISSN: 2307-1354.2017.–С.19-24.
6. Ильин Д.В., Зайцева Т.С. «Помехоустойчивое кодирование» – Научно-информационный издательский центр "Институт стратегических исследований" (Москва). ISSN: 2414-9918.2020. – С.512-518.
7. Чигликова Н.Д., Димов Д.Е. «Разработка алгоритма визуализации кода Хемминга и обратное преобразование» – Кубанский государственный технологический университет (Краснодар). – С. 469-476.

A.P. Zaitsev, A.V. Semashko

ELIMINATION OF INVERSION IN CYCLIC CODES TO IMPROVE THE NOISE IMMUNITY OF COMMUNICATION SYSTEMS

Nizhny Novgorod State Technical University n. a. R. E. Alekseev

The method of binary noise-proof coding with the possibility of detecting the inversion of all characters in the code block is considered. By means of simulation modeling, a study of codes based on cyclic codes is carried out, which allow, in addition to single errors, to detect and correct the inversion of the characters of the code block. The Hamming codes (15, 11) and the BCH code (31, 16) were studied in the Matlab software environment.

Keywords: modeling, cyclic codes, noise-proof coding, communication systems, control system.

МОДЕЛИРОВАНИЕ КОРРЕЛЯЦИОННЫХ ХАРАКТЕРИСТИК ПСЕВДОСЛУЧАЙНЫХ ПОСЛЕДОВАТЕЛЬНОСТЕЙ ДЛЯ РАЗЛИЧЕНИЯ СИГНАЛОВ

Нижегородский государственный технический университет им. Р.Е. Алексеева

Обобщены существующие методики анализа корреляционных характеристик псевдослучайных характеристик для различения сигналов в программной среде Matlab.

Ключевые слова: моделирование, псевдослучайные последовательности, функция неопределенности, системы связи, автокорреляция.

На выходе согласованного фильтра приемника радиосигнала формируется взаимно корреляционная функция опорного сигнала и наблюдаемого процесса. Если в наблюдаемом процессе присутствует опорный сигнал, то на выходе согласованного фильтра формируется автокорреляционная функция (АКФ) полезного сигнала, к которой добавлен шум канала связи, прошедший через согласованный фильтр. В радиолокации и связи часто возникает проблема различения сигналов одинаковой структуры. В радиолокации для различения близко идущих целей, в системах связи для борьбы с многолучевым распространением сигналов.

Возможным способом решением вышеизложенных проблем является расширение частотного спектра сигналов. Широко известными методами формирования сигналов с большой базой [1] – формирование сигналов с частотной внутриимпульсной модуляцией или формирование сигналов с фазокодовой манипуляцией. Фазоманипулированный сигнал – это совокупность n сомкнутых прямоугольни парциальных радиоимпульсов одинаковой длительности и частоты при ограниченном числе p различающихся возможных значений начальной фазы колебаний.

Для фазовой модуляции используются различные псевдослучайные наборы, в том числе коды Баркера, М-последовательности и другие. Коды Баркера являются наиболее эффективными последовательностями в соотношении центрального и боковых лепестков АКФ в зависимости от длины кода [2]. Однако коды Баркера ограничены максимальной длиной в 13 символов что весьма ограничивает потенциальные возможности применения данных кодовых последовательностей. М последовательности имеют более худшее отношение высоты основного лепестка к боковым АКФ, но не ограничены длиной последовательности. М-последовательности или последовательности максимальной длины могут быть описаны с помощью рекуррентных формул. Основными свойствами являются

1. М-последовательности являются периодическими с периодом $N = 2^n - 1$
2. Количество символов, принимающих значение единица, на длине одного периода М-последовательности на единицу больше, чем количество символов, принимающих значение нуль
3. Любые комбинации символов длины n на длине одного периода М-последовательности за исключением комбинации из n нулей встречаются не более одного раза. Комбинация из n нулей является запрещённой: на её основе может генерироваться только последовательность из одних нулей
4. Сумма по модулю 2 любой М-последовательности с её произвольным циклическим сдвигом также является М-последовательностью
5. Периодическая АКФ любой М-последовательности имеет уровень боковых лепестков, пропорциональный $\frac{1}{N}$.

6. АКФ усечённой М-последовательности, под которой понимается непериодическая последовательность длиной в период N , имеет величину боковых лепестков, близкую к $-\frac{1}{\sqrt{N}}$. Поэтому с ростом N величина боковых пиков уменьшается.

На основе М-последовательностей можно получить большой ансамбль квазиортогональных двоичных секвенций с хорошими корреляционными свойствами, число которых во много раз превосходит количество исходных наборов. Таковыми являются коды Голда и Кассами.

Псевдослучайная последовательность Голда не ограничена по длительности и формируется при помощи подходящих пар кодов максимальной длины, которые складываются по модулю два. После преобразования двух наборов получается один с длиной равной М-последовательности, но количеством различных версий кода одной длины равно $N = L+2$ где, L длина псевдослучайного ряда [4].

Коды Кассами также формируются на основе М-последовательностей. Генерируются они при помощи процедуры децимации, комбинированной обработки наборов Голда и уже сформированных последовательностей Кассами. Они представляют интерес из-за того, что их взаимно корреляционная функция для различных наборов различна.

Приведенная ниже процедура моделирования предназначена для того, чтобы получать значения автокорреляционных функций исследуемых последовательностей. Для реализации был выбран пакет приложения MATLAB. Вычисление значений было выполнено при помощи встроенных стандартных библиотек. Они позволяют работать эффективно с большими массивами значений и выводить результат в наглядной форме при помощи графиков, которые легко настраиваются [5].

Рассмотрим функцию неопределенности [6] исследуемых кодов. Она содержит исчерпывающую информацию о автокорреляционных свойствах соответствующих сигналов.

Графики на рис. 1 функции неопределенности кода Баркера длиной 13 при сечении при нулевом отклонении по частоте боковые лепестки начинаются на уровне $1/13$ от основного, в сечении при нулевом отклонении по времени боковые лепестки в среднем находятся на уровне 0,1 от высоты основного лепестка на всей полосе частот сигнала.

Характеристики рассматриваемых в дальнейшем сигналов были – частота 20 МГц и длительность 3 мс. Следующие для рассмотрения были взяты графики функции неопределенности кода Кассами длиной 63.

Из графиков на рис. 2 видно, что АКФ имеет боковые лепестки, находящиеся на уровне $1/5$ от основного лепестка и уменьшаются при увеличении разницы во времени. Также можно увидеть, что в сечении при нулевом отклонении во времени значения функции при отклонении на 0,2 кГц становятся незначительной что положительно влияет на соседний сигнал по частоте.

Графики на рис. 3 функции неопределенности кода Голда длиной 63 при сечении при нулевом отклонении по частоте боковые лепестки начинаются на уровне 0,5 от основного после чего спадают до нуля, в сечении при нулевом отклонении по времени боковые лепестки в среднем находятся на уровне 0,25 от высоты основного лепестка на всей полосе частот сигнала.

Последним для рассмотрения были взяты графики функции неопределённости М-последовательности длиной 63 на рис. 4. Боковые лепестки при сечении при нулевом отклонении по частоте начинаются на уровне 0,05 от основного, после чего увеличиваются до 0,1 и в среднем остаются на всей длительности сигнала. В сечении при нулевом отклонении во времени наблюдается похожая ситуация, как и с кодом Кассами, максимум боковых лепестков на уровне 0,2 от основного, но при отклонении на 2 кГц становится незначительной.

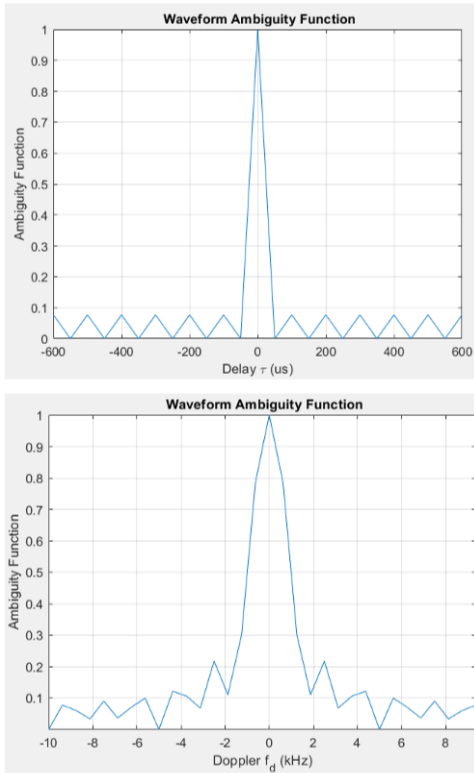


Рис. 1. Графики функции неопределенности кода Баркера длиной 13 в сечении, сверху при нулевом отклонении по частоте, снизу при нулевом отклонении во времени

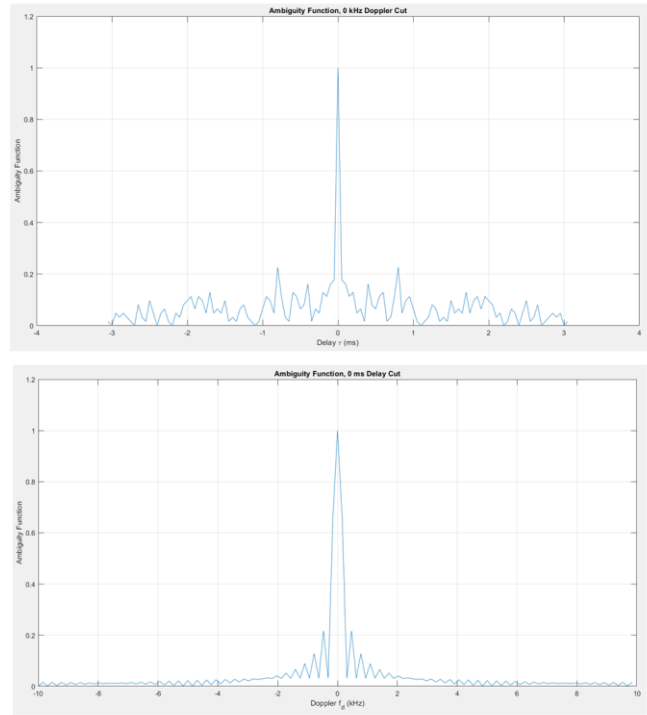


Рис. 2. Графики функции неопределенности кода Кассами длиной 63 в сечении, сверху при нулевом отклонении по частоте, снизу при нулевом отклонении во времени

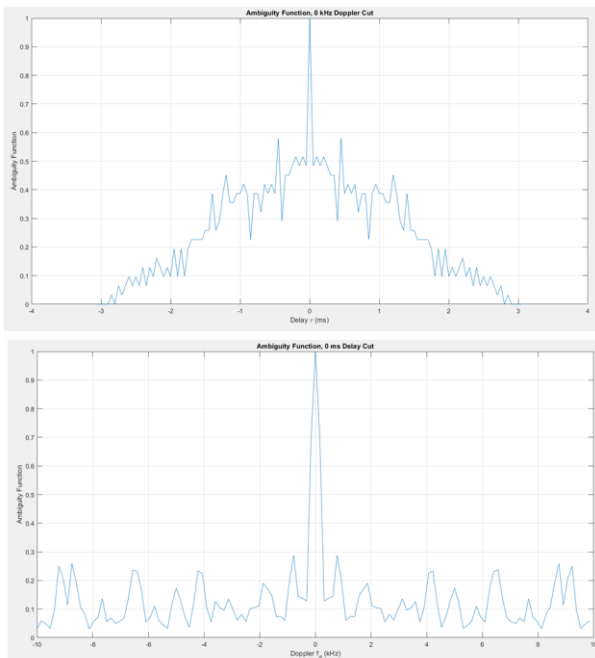


Рис. 3. Графики функции неопределенности кода Голда длиной 63 в сечении, сверху при нулевом отклонении по частоте, снизу при нулевом отклонении во времени

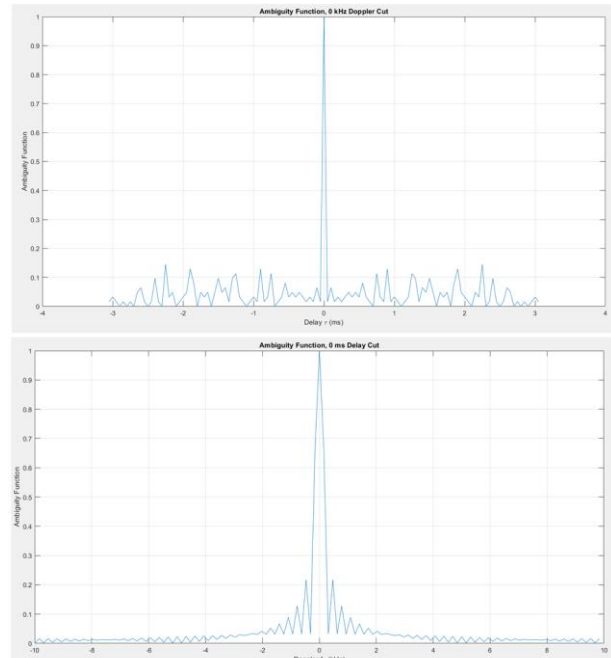


Рис. 4. Графики функции неопределенности M-последовательности длиной 63 в сечении, сверху при нулевом отклонении по частоте, снизу при нулевом отклонении во времени

Как видно из графиков существенные различия имеют функция неопределенности кода Голда и Кассами с М-последовательностью. Между последними различий. Коды Голда существенно уступают по значениям в сечении при нулевом отклонении по частоте, и в меньшей степени уступают в сечении при нулевом отклонении по времени.

Как можно увидеть, из результатов все представленные шумоподобные сигналы уступают кодам Баркера по отношению величины основного лепестка к боковым в зависимости от длины кода, однако иные шумоподобные сигналы не имеют жестких ограничений по длительности, что служит основанием для их эффективного использования.

Библиографический список

1. Анализ кодовых последовательностей для построения асинхронных систем передачи данных И.В. Филиппенко 6/3 (42) 2009
2. Ипатов В.П. Широкополосные системы и кодовое разделение сигналов. – М: Мир связи, 2007. – 488 с.
3. Баринов В.В., Лебедев М.В., Кузнецов В.С. Анализ корреляционных характеристик расширяющих ансамблей // Электросвязь. 2006. № 3. С. 38-39
4. Генерация ансамблей кодов Голда для систем прямого расширения спектра Кузнецов В.С., Шевченко И.В., Волков А.С., Солодков А.В. 2017. № 96
5. Radar systems & analysis and design using Matlab. Mahafza, Bassem R.
6. Тузов Г. И. Статистическая теория приема сложных сигналов. М., «Сов радио», 1977, с. 400

R.R. Meshеров, A.V. Semashko

MODELING THE CORRELATION CHARACTERISTICS OF PSEUDO RANDOM SEQUENCES FOR DIFFERENTIATING SIGNALS

Nizhny Novgorod State Technical University named after R.E. Alekseev

The existing techniques for analyzing the correlation characteristics of pseudo-random characteristics for distinguishing signals in the Matlab software environment are generalized.

Keywords: modeling, pseudo-random sequences, uncertainty function, communication systems, autocorrelation.

РАЗРАБОТКА ПРОЦЕДУРЫ ВИДЕОАНАЛИТИКИ ДЛЯ ОБЕСПЕЧЕНИЯ ОБЩЕСТВЕННОЙ БЕЗОПАСНОСТИ

Нижегородский государственный технический университет им. Р.Е. Алексеева

В статье описано проектирование системы обработки и хранения информации с камер видеонаблюдения в долгосрочной перспективе, для восстановления траекторизации перемещений определенного человека в определенный период времени.

Ключевые слова: распознавание лиц, видеоаналитика, хранение данных, нейронные сети, алгоритмы распознавания.

Введение

В настоящее время практически каждый двор, крупный перекресток, общественное место в условиях большого города оборудованы камерами наблюдения. Запись ведется непрерывно, а в объектив камеры ежедневно попадает тысячи людей. Что ведет к монотонному возрастанию объемов занимаемой памяти на серверах. Большинство камер хранят данные 5-30 дней, и в случае, если информация за данный период оказалась не востребована, информация уничтожается [10]. Но так как срок давности в уголовном праве исчисляется сроком от 2 до 15 лет [3], встает вопрос о хранении некоторой информации в течение длительного периода времени. Таким образом, становится задача обработки и хранения информации с камер видеонаблюдения в долгосрочной перспективе, для восстановления траекторизации перемещений определенного человека в определенный период времени.

Эволюция технологий распознавания

Первые упоминания распознавания лиц можно отнести к XIX веку. Для противодействия преступности, полиция некоторых Американских городов распространяла объявления с фотографиями, описанием преступлений и сообщением о розыске преступников. Объявления получали распространения в местах массового скопления людей, а также в почтовых отделениях.

Сегодня важную роль в процессах распознавания лица человека играют современные технологии. Такие как социальные сети, интернет, телевидение и другие средства связи. Развитие искусственного интеллекта, а также биометрических технологий привело к тому, что повысилась доступность компьютеризованных технологий распознавания лиц. Сегодня распознавание может быть проведено в большом масштабе, а также в сложных условиях [1].

В 60-е годы XX века разработали первую компьютеризированную систему распознавания лиц. Первым разработал данную систему американский ученый в области искусственного интеллекта Вудро Уилсон Бледсоу [2]. В своей работе он использовал способ классификации лиц с помощью линий сетки. Но в данной технологии всё еще требовалось участие человека. Он вручную отмечал отличительные черты лица на фотоснимках и фиксировал данную информацию в компьютер. С данной технологией скорость проверки достигала 40 лиц в час, что считалось очень высоким результатом [5]. К концу 60-х годов технология развивалась в научно-исследовательском институте Стэнфорда. Точность распознавания лиц у системы оказалась лучше, чем у людей. В 90-х годах самое эффективное решение было реализовано в Рурском университете, в Германии. Новая технология имела очень высокую точность. Считается, что с этого момента произошел рывок технологий распознавания лиц [6].

Принцип работы

Распознавание лиц – биометрическая, бесконтактная система идентификации человека по лицу.

Процесс сопоставления лиц, попавших в объектив камеры, с эталоном, занесенном в БД. На данный момент существует три распространенные схемы реализации системы:

- Анализ видеопотока на сервере

Наиболее распространенная схема реализации – камера передает видеопоток на сервер, где ПО выполняет анализ видеопотока и сравнивает полученные изображения лиц, с БД эталонов.



Рис. 1. Анализ видеопотока на сервере

- Анализ видеопотока на IP-камере

При данной реализации, видеопоток анализируется на самой камере, а на сервер передаются метаданные.



Рис. 2. Анализ видеопотока на IP-камере

- Анализ видеопотока на устройстве контроля доступа

В отличие от первых двух систем, в текущей системе камера встроена в устройство контроля доступа (УКД), в котором помимо распознавания лица, происходящем на устройстве, выполняется функция управления доступом. База данных эталонных лиц хранится на УКД [7]/



Рис. 3. Анализ видеопотока на устройстве контроля доступа

Таблица 1

Преимущества и недостатки схем реализации распознавания

	Преимущества	Недостатки
Анализ видеопотока на сервере	Возможность использования уже существующей системы	Высокая нагрузка на сеть, высокая стоимость сервера.

	видеонаблюдения.	
Анализ видеопотока на IP-камере	Подключение неограниченного количества камер к одному серверу.	Необходимость наличия специальных камер. Высокая стоимость. Взаимодействие софта на камере и на сервере. Размер БД эталонов.
Анализ видеопотока на устройстве контроля доступа.	Низкая стоимость систем.	Устройства выпускаются для использования в помещениях.

Безопасный город

«Безопасный город» - комплекс программно-аппаратных средств и организационных мер для обеспечения видеоохраны и технической безопасности, а также управления в едином информационном пространстве объектами жилищно-коммунального хозяйства и другими распределенными объектами.

Предпосылки создания АПК «Безопасный город»:

- Низкий уровень безопасности населения;
- Низкий уровень безопасности инфраструктурных объектов;
- Большое время реагирования экстренных служб;
- Высокая аварийность на автомобильных дорогах;
- Риск возникновения чрезвычайных ситуаций;
- Отсутствие централизованной системы оповещения [9].

В рамках проекта «безопасный город», перед системой ставят следующие задачи:

- Обеспечение безопасности жителей города в общественных местах, жилых объектах и транспорте;
- Оказание помощи в работе правоохранительных структур в обеспечении правопорядка;
- Контроль предприятий, выбрасывающих химически опасные вещества;
- Координирование работы снегоуборочной технике;
- Предотвращение беспорядков во время массовых мероприятий [4].

Проектирование системы

В данной системе планируется реализовать три основных модуля:

- Модуль распознавания лица в видеопотоке;

Данный модуль включает в себя алгоритмы распознавания лица, в непрерывном видеопотоке. Анализирует входные данные с определенной камеры, сравнивает данные с эталонными и передает информацию на обработку следующему модулю.

- Модуль идентификации и обработки информации в базе данных;

Данный модуль проводит анализ информации, принятой от модуля распознавания.

Заносит данные в общую БД в следующем виде:

23d1ac51ac 05.06.20 12:20:21 111.123456 111.123456 –48 байт

где

- 23d1ac51ac -ID человека – 10 символов, в шестнадцатеричном виде
- 05.06.20 - дата
- 12:20:21 - время попадания на камеру
- 111.123456 111.123456 - координаты.

Параллельно система комплексирования определяет, есть ли записи с одинаковым IDи попадающие на разные камеры, в радиусе не более 500 метров, с разницей во времени не более 20 минут. В случае нахождения таких записей, система в результирующей БД заносит точку нахождения объекта в определенное время, в определенном месте.

Хранится информация в виде, представленном выше. Таким образом, одна запись занимает всего 48 байт. В случае, если определенный человек попадает на камеру каждую минуту в течение дня, в хранилище данных информация займет примерно 70 Килобайт.

Соответственно, в идеальных условиях, за год на одного человека потребуется 25.55 Мегабайт. А для хранения информации о перемещении целого города с населением в 1 миллион человек необходимо примерно 25 терабайт.

– Модуль отображения траектории на карте.

По запросу, по определенному человеку, указав дату, система отрисует на карте, по координатам перемещение человека в течении дня, с указанием времени нахождения на определенном участке.

Заключение

Таким образом, спроектирована система, анализирующая непрерывный видеопоток с камеры видеонаблюдения, анализирующий его, и передающий на хранение на длительный период времени. Благодаря спроектированной системе, данные о перемещениях человека можно хранить несколько десятилетий, при этом занимать пространство на жестком диске не более 26 мегабайт в год.

Библиографический список

1. Брилюк Д. Распознавание человека по изображению лица нейросетевыми методами / Д. Брилюк, В. Старовойтов. – Минск 2002
2. Boyer A. Automated reasoning: Essays in Honor of Woody Bledsoe / A. Boyer, R. Boyer. – Kluwer academic publishers group 1991
3. Конституция Российской Федерации (принята всенародным голосованием 12.12.1993) // Собрание законодательства РФ.
4. Белоусов, А. «Умный город» как основа цифровой экономики / А. Белоусов. – Москва 2017
5. URL: <http://wired.com/story/secret-history-facial-recognition/>
6. URL: http://greenvision.ua/blog/articles/Tekhnologiya_raspoznavaniya_list
7. URL: <http://securityrussia.com/blog/face-recognition.html>
8. URL: <https://22.xn-b1aew.xn--p1ai>
9. URL: https://dssl.ru/integratoram/solution/safe_city.php
10. URL: <https://video.dit.mos.ru/faq/#:~:text=Срок%20хранения%20архивной%20видеоинформации%20с,объектов%20строительства%20%2030%20календарных%20дней>

A.A. Muravov, A.V. Semashko, O.E. Vinogradov

DEVELOPMENT OF A VIDEO ANALYTICS PROCEDURE FOR COUNTERING TERRORISM

Nizhny Novgorod State Technical University named after R. E. Alekseev

The article has designed a system for processing and storing information from video surveillance cameras in the long term, to restore the trajectory of movements of a certain person in a certain period of time.

Keywords: face recognition, video analytics, data storage, neural networks, recognition algorithms.

АДАПТАЦИОННЫЕ МЕХАНИЗМЫ ДЛЯ ПОВЫШЕНИЯ НАДЕЖНОСТИ АВИАЦИОННОЙ ЗАГОРИЗОНТНОЙ ДКМВ СВЯЗИ

*Нижегородский государственный технический университет им. Р.Е. Алексева,
ООО «НПП Прима», г. Нижний Новгород*

В работе приведены основные подходы, лежащие в основе построения современных авиационных загоризонтных декаметровых систем связи. Рассмотрены адаптационные механизмы разных уровней модели OSI, показана целесообразность их применения. Рассмотрены особенности ДКМВ канала, и даны оценки повышения качества связи.

Ключевые слова: авиационная связь, адаптивные системы связи, ионосферный ДКМВ канал, надежность связи, оценка качества канала.

Загоризонтная ДКМВ связь в настоящее время является основным средством передачи речи и данных на радиоперелиниях магистральных самолетов и диспетчеров управления воздушного движения гражданской авиации России и других стран ИКАО. Применение адаптивных алгоритмов и устройств позволяет значительно повысить эффективность и надежность радиосвязи. ДКМВ связь позволяет производить передачу данных на дальности до нескольких тысяч километров, при относительно малой мощности передающей аппаратуры и без использования ретрансляции. При этом, в отличие от, например, спутниковой связи ДКМВ связь является экономически более выгодной и более простой в развертывании. Для нашей страны, обладающей огромной территорией со значительно высокой неравномерностью урбанизации и плотностью населения, а также в условиях широкого разброса по климатическим зонам, развитие ДКМВ связи остается одной из приоритетных задач.

ДКМВ канал является одним из наиболее сложным в плане ведения связи. Для ионосферного канала характерны нестационарные замирания, обусловленные многолучевостью, а также высокий уровень шумов. Кроме того, для авиационной связи характерны существенные динамически меняющиеся частотные расстройки, обуславливающиеся эффектом Доплера и помехи иных станций, при использовании общего ресурса частотного канала (системы TDMA).

Для повышения эффективности ведения связи в таких условиях, в настоящее время разрабатываются и разрабатываются современные и перспективные адаптивные системы. Такие системы используют несколько уровней адаптации, которые подразумевают как административные действия, так и использование высокоэффективных алгоритмов, для повышения качества связи.

Выделяют три уровня адаптивных систем связи:

Первый уровень затрагивает вопрос выделения рабочих частот на стадии формирования плана связи. Данный этап подразумевает анализ метеоданных, а также статистики состояния ионосферы для конкретной локации развертывания системы связи таким образом, чтобы задействовать те частоты, которые в момент предполагаемого функционирования имеют наилучшее прохождение сигнала. На данном уровне активно используется статистика прохождения сигналов и чем она более полна, тем лучше будет функционирование системы.

Второй уровень адаптации подразумевает под собой работу радиоаппаратуры в период проведения сеансов связи. Для повышения эффективности связи, подразумевающей тут максимизацию эффективной скорости передачи данных при обеспечении минимально приемлемой вероятности ошибки, аппаратура связи должна

обеспечивать работу в различных режимах, использующих различные сигнально-кодовые конструкции. Понятно, что возможно добиться повышение помехоустойчивости путем понижения скорости передачи данных. Это может производиться как за счет увеличения длительности передаваемых сигнальных символов, так и, что более предпочтительно, введением помехоустойчивого кодирования с различной мощностью кода. Кроме того, использование разных видов модуляции повышает общую эффективность, так как различные сигналы обладают различной стойкостью к разным видам помех (частотные расстройки, АГБШ, замирения, частотно-сосредоточенные помехи).

Таким образом, для повышения качества связи, приемная аппаратура должна обеспечивать слежение за состоянием канала, оценивать его качество и сохранять статистику «слышимости» всех абонентов. При необходимости установления соединения с конкретным абонентом, аппаратура должна проанализировав наколенную статистику, выбрать тот режим работы, который обеспечит максимальную скорость передачи данных, обеспечив при этом минимально допустимую вероятность ошибки.

Третий уровень адаптации подразумевает использование алгоритмов приемной аппаратуры, которые позволяют обеспечить наилучший прием сообщений в условиях ДКМВ канала. Такие алгоритмы должны в течении принимаемого сообщения производить оценку качества канала, и подстройку демодулятора для лучшего приема в условиях изменяющейся расстройки частоты (эффект Доплера) и замираний (эффекты многолучевости). Данные алгоритмы должны оценивать допустимый интервал стационарности и производить оперативную оценку канала (например, импульсная характеристика), чтобы минимизировать ошибку демодуляции данных. Кроме того, именно на этом уровне производится измерение SINAD и иных параметров, которые в дальнейшем используются на уровне 2 при установлении наиболее оптимального режима передачи данных.

В своей работе мы разработали алгоритмы адаптации третьего (канального) уровня и промоделировали их работу в условиях ДКМВ канала. На рисунках 1, 2 и 3 показаны сравнительные графики помехоустойчивости без использования разработанных алгоритмов и с ними. Сравнение проводилось при прочих равных (вид излучения J3E, информационная полоса 3 кГц, модуляция PSK, скорость манипуляции 2400 бод, помехоустойчивое сверточное кодирование, перемежитель на 4.8 сек).

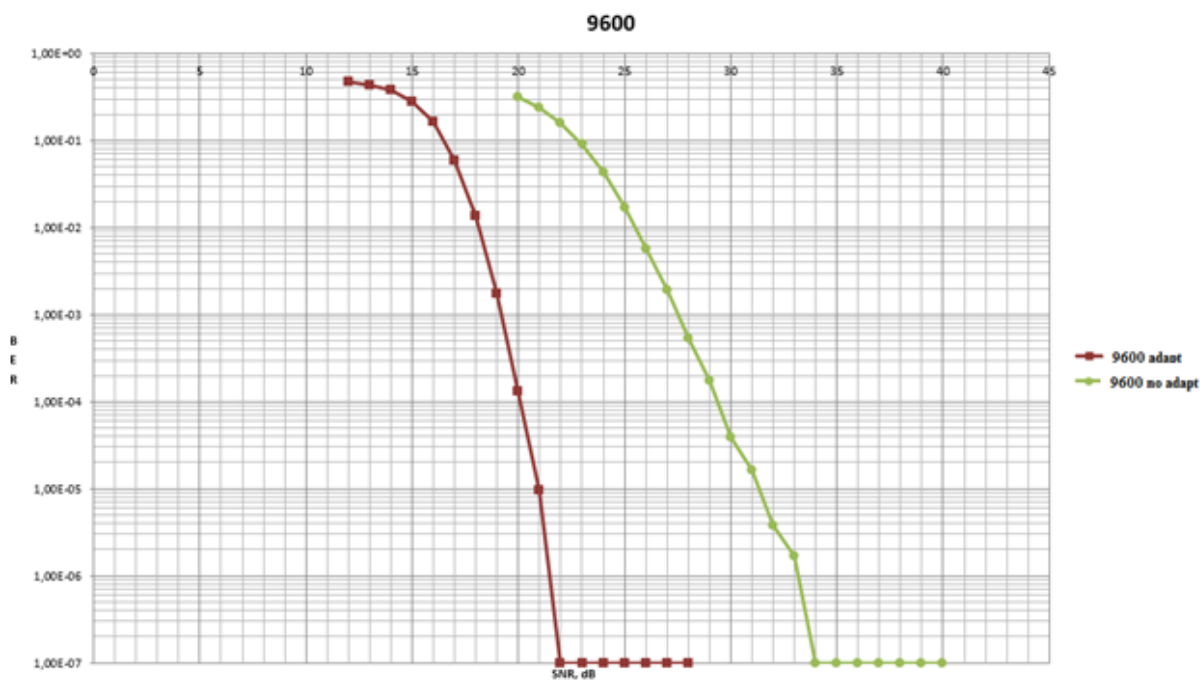


Рис. 1. Сравнение помехоустойчивости для эффективной скорости 9600 бит/с

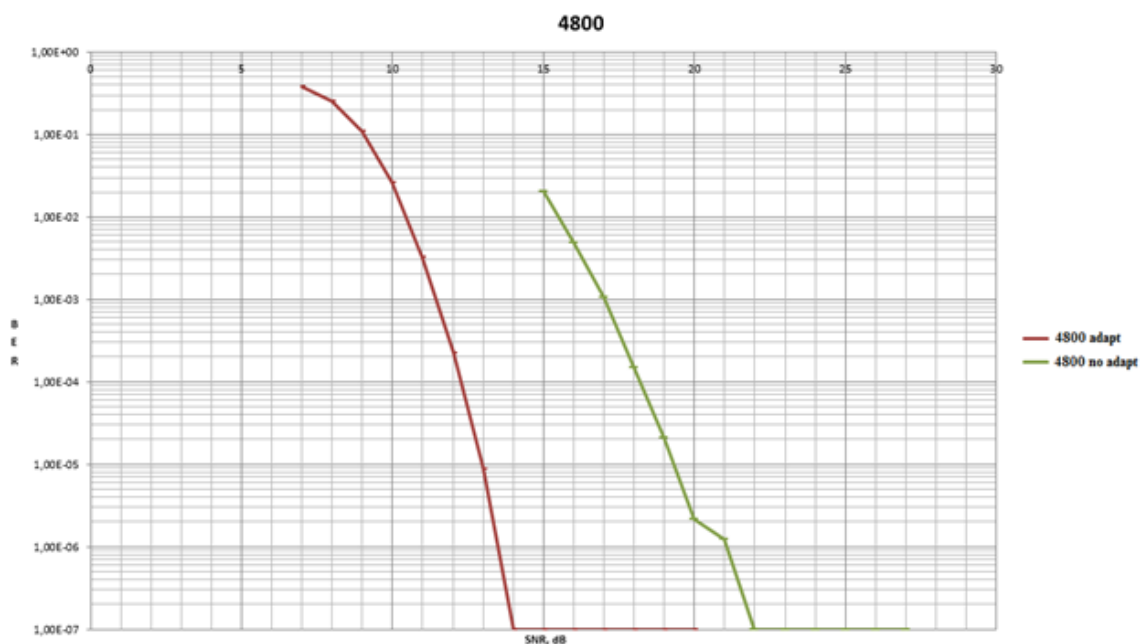


Рис. 2. Сравнение помехоустойчивости для эффективной скорости 4800 бит/с

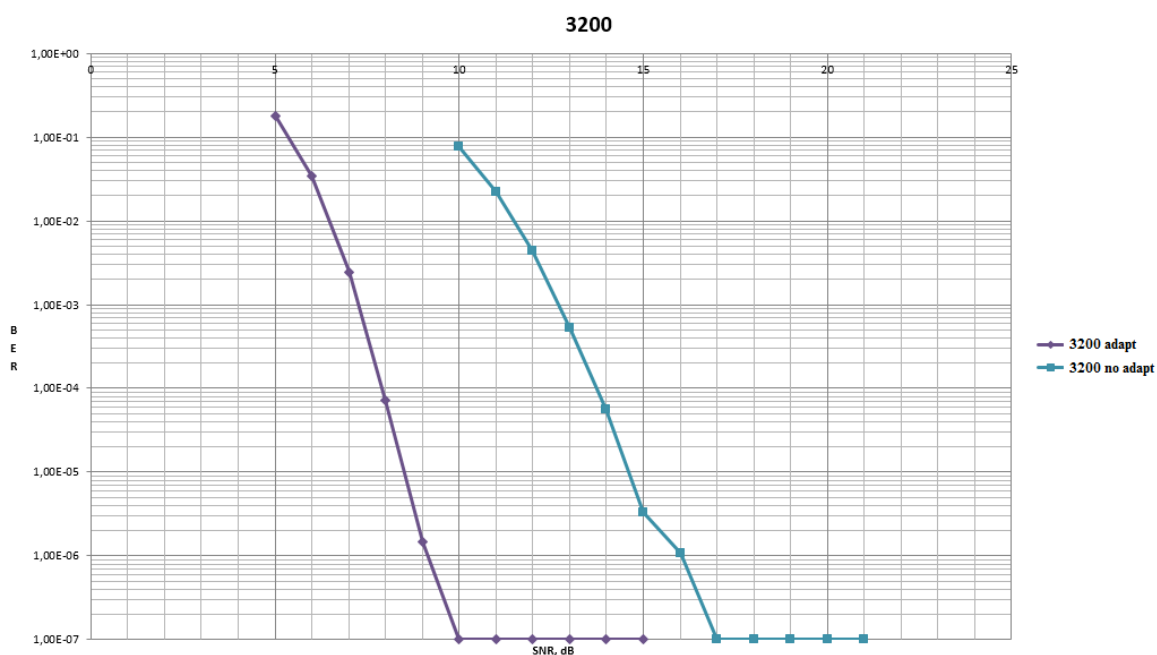


Рис. 3. Сравнение помехоустойчивости для эффективной скорости 3200 бит/с

Таким образом, для успешного функционирования современной адаптивной авиационной связи ДКМВ диапазона необходимо выполнения процедур на нескольких уровнях:

- производить выбор рабочих частот согласно статистике прохождения радиосигнала, для предполагаемых условий функционирования системы;
- производить выбор оптимального режима работы передачи данных для проведения сеанса связи. Выбор должен основываться на статистике прослушивания абонентов и составления оценок качества канала между ними в ходе текущего прослушивания эфира;
- производить текущую оценку изменяющегося канала во время приема сигнала и производить оперативную подстройку приемной аппаратуры.

Библиографический список

1. MIL-STD-188-141B. Interoperability and performance standards for medium and high frequency radio systems. DOD interface standard. 1999.
2. ARINC Specification 635-2. HF Data Link Protocols. 1998.
3. MIL-STD-188-110C. Interoperability and Performance Standards for Data Modems. 2011. 87
4. Lim D.W., Heo S.J., No J.S. An overview of peak-to-average power ratio reduction schemes for OFDM signals // Journal of communications and networks. 2009. Vol. 11. No. 3. P. 229-239.
5. Бочков Г.Н., Горохов К.В., Колобков А.В. Алгоритмы демодуляции OFDM-сигналов в каналах с рассеянием во времени и по частоте // Известия вузов. Радиофизика. 2016. Т. 59. № 1. С 67–89.
6. Лосев Ю.И., Бердников А.Г. и др. Адаптивная компенсация помех в каналах связи / Под ред. Ю.И. Лосева. М.: Радио и связь, 1988. 208 с.

A.A. Yampolsky, A.V. Semashko

ADAPTIVE ALGORITHMS FOR IMPROVING THE RELIABILITY OF THE AVIATION OVER-THE-HORIZON HF COMMUNICATION SYSTEMS

*Nizhny Novgorod State Technical University n. a. R.E. Alekseev,
PRIMA, Nizhny Novgorod*

The article presents the main approaches underlying the construction of modern aviation over-the-horizon HF communication systems. The adaptive mechanisms of different levels of the OSI model are considered, and the expediency of their application is shown. Considered the features of the HF channel, and shown the improvement quality of the communications.

Keywords: aviation communication systems, adaptive communication systems, ionospheric HF channel, communication reliability, estimation of HF channel quality.

ОПТИМИЗАЦИЯ ПОСТРОЕНИЯ КАМПУСНЫХ СЕТЕЙ

Нижегородский государственный технический университет им. Р.Е. Алексеева

В работе описаны основные принципы построения кампусных сетей приведен детальный анализ структуры сетей кампусов. Рассмотрена трех уровневая модель сети, состоящая из составляющих: уровня ядра обеспечивающего связь с «внешним» миром, т.е. выход во всемирную паутину; уровня распределения – обеспечивающего маршрутизацию внутри кампуса между локальными сетями и доступ к сетевым сервисам организации; уровня доступа – обеспечивающего доступ техники к сети на территории кампуса.

Ключевые слова: локальная сеть, кампусная сеть, архитектура построения сети, доступ, ресурсы, структура сети.

Локальные сети, объединяющие конечных пользователей, включаются в состав корпоративной сети. Названия структурных единиц корпоративной сети отражают не только территорию покрытия, но и организационную структуру предприятия. Так, принято делить корпоративную сеть на сети отделов и рабочих групп, сети зданий, кампусные сети, магистраль.

В настоящее время основными коммуникационными устройствами кампусных сетей является коммутаторы, а не маршрутизаторы или концентраторы. Это обуславливается несколькими причинами. Во-первых, стоимость коммутаторов в последнее время значительно снизилась и приблизилась к стоимости концентраторов. Так как полная пропускная способность коммутатора значительно больше, чем у концентратора, при одинаковой стоимости предпочтение отдается коммутаторам. Во-вторых, в качестве магистральных коммутационных устройств в настоящее время применяются коммутаторы третьего уровня, которые позволяют фильтровать широкополосный трафик, и в тоже время обладают большим быстродействием, чем маршрутизаторы. Показатель «цена/ производительность» рассчитанный для одного порта у таких коммутаторов ниже, чем аналогичный показатель маршрутизаторов. Многие проектировщики сетей придерживаются принципа «коммутаторы – где возможно, маршрутизаторы – где необходимо».

Распространённым вариантом архитектуры построения сетей является трехуровневая модель, предложенная компанией Cisco. В данную модель входят следующие уровни:

- Уровень доступа (AccessLevel) - на данном уровне организуется доступ пользователей к ресурсам сети;
- Уровень распределения (DistributionLevel) или агрегации - на данном уровне организуется агрегация потоков данных пользователей и распределения этих потоков по каналам связи в горизонтальной плоскости уровня или между ниже- и вышележащими уровнями;
- Уровень ядра (CoreLevel) – на данном уровне организуется высокоскоростная магистраль передачи данных между сетями, а также доступ к высоконагруженным ресурсам, которые могут находиться на Серверной ферме предприятия или иного провайдера услуг.

В реальности существуют как варианты сетей, соответствующие данной модели, так и варианты сетей, в которых один уровень может совмещать функционал двух сетей одновременно.

Виды сетей передачи данных

Сети отделов — это сети, которые используются сравнительно небольшой группой сотрудников, работающих в одном отделе предприятия. Эти сотрудники решают некоторые общие задачи, например, ведут бухгалтерский учет или занимаются маркетингом. Считается, что отдел может насчитывать до 100-150 сотрудников.

Сеть отдела – это локальная сеть, которая охватывает все помещения, принадлежащие отделу. Это могут быть несколько комнат или этаж здания.

Главным назначением сети отдела является разделение локальных ресурсов, таких как приложения, данные, лазерные принтеры и модемы. Обычно сети отделов не делят на подсети и в их состав входят один или два файловых сервера. В этих сетях локализуется большая часть трафика предприятия.

Задачи сетевого администрирования на уровне отдела относительно просты, поэтому на небольших предприятиях эти задачи может выполнять один из сотрудников отдела, который не входит в состав персонала администрирования сетей передачи данных, но обладающий достаточным уровнем знаний в этой области.

Задачи сетевого администрирования на уровне отдела следующие:

- добавление новых пользователей;
- устранение простых отказов;
- установка новых узлов и новых версий программного обеспечения.

Сети рабочих групп

Существует и другой тип сетей, близкий к сетям отделов, – сети рабочих групп. К таким сетям относят совсем небольшие сети, включающие до 10-20 компьютеров. Характеристики сетей рабочих групп практически не отличаются от описанных характеристик сетей отделов. Такие свойства, как простота сети и однородность, здесь проявляются в наибольшей степени, в то время как сети отделов могут приближаться в некоторых случаях к следующему по масштабу типу сетей – сетям зданий и кампусов.

Кампусная сеть

Частным случаем корпоративных сетей являются кампусные сети, то есть вычислительные сети, объединяющие все вычислительные устройства в одном здании, или нескольких близлежащих зданиях. В настоящее время кампусные сети – один из наиболее распространенных классов вычислительных сетей, кампусная сеть есть практически в любой крупной организации. Корпоративная сеть крупного предприятия может объединять несколько кампусных сетей, находящихся в различных городах или странах. Кампусные сети обладают определенными особенностями, налагающими соответствующие требования на проектирование. К числу таких особенностей относятся:

1. Относительно небольшая площадь, и, как следствие, использование для передачи данных высокоскоростных технологий локальных сетей.
2. Использование в качестве основного коммуникационного устройства не маршрутизаторов и концентраторов, а коммутаторов различной мощности.
4. Использование структурированных кабельных систем.
5. Гетерогенность сети (одна кампусная сеть может состоять из подсетей, построенных на базе различных сетевых технологий).

Каждую корпоративную сеть можно рассматривать как набор локальных сетей и глобальных сетей, которые соединяют LAN между собой. Кампусная сеть является одной из составляющих корпоративной сети. Она покрывает одно или несколько зданий, находящихся рядом друг с другом. Для проектировки используется трех уровневая модель (рис. 1), которая состоит из следующих составляющих:

- Уровень ядра – обеспечивает связь с «внешним» миром, т.е. выход во всемирную паутину;
- Уровень распределения – обеспечивает маршрутизацию внутри кампуса между локальными сетями и доступ к сетевым сервисам организации;
- Уровень доступа – обеспечивает доступ техники к сети на территории кампуса. Каждый из уровней выполняет свои функции. Они будут рассмотрены чуть позже.

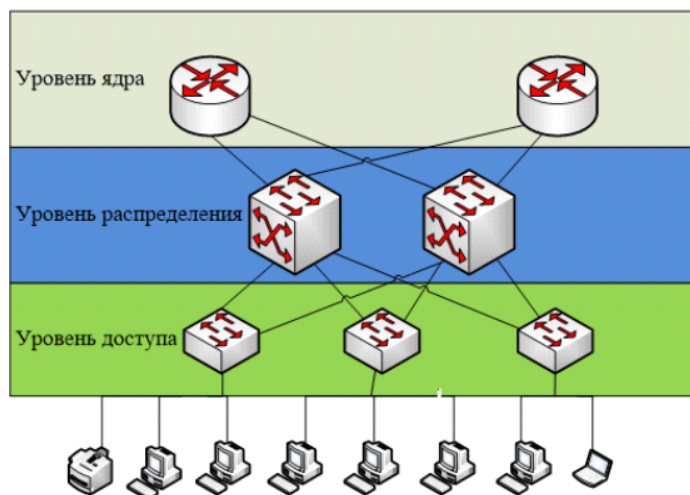


Рис. 1. Иерархия сети

Сравнивая иерархическую сеть с плоской, можно выделить ряд преимуществ у иерархической. Во-первых, это локализация трафика, т.е. между уровнями передаются только те данные, которые передаются между уровнями. Во-вторых, в иерархической сети трафик может быть отфильтрован на втором уровне, что значительно повышает уровень безопасности всей сети. В-третьих, в иерархической сети лучше отлажено управление широковещательными потоками трафика, что позволяет значительно снизить нагрузку на сетевые устройства и повысить скорость трафика в сети. В каждом отдельном здании или нескольких зданиях, расположенных рядом друг с другом организуется кампусная сеть. Соединив их между собой, например, по VPN можно получить полноценную корпоративную сеть, как на рис. 2:

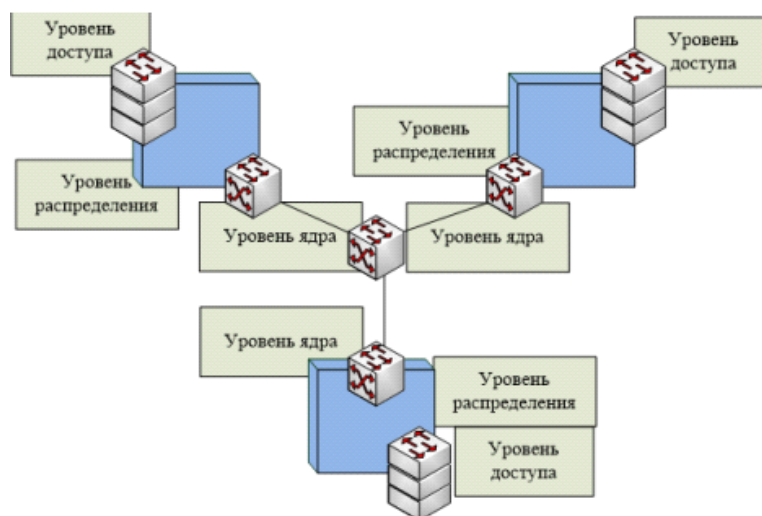


Рис. 2. Объединение нескольких кампусных сетей в единую корпоративную сеть

Ещё одним немаловажным фактом в проектировки кампусной сети, да и не только кампусной, является отсутствие какой-либо привязки к определённой схеме. В разных ситуациях приходится применять самые разные структуры сети, иногда даже объединять. Та же иерархическая схема сети имеет два вида: первый – это трёхуровневая модель, которая было рассмотрена выше, вторая – это двухуровневая модель, в ней уровень ядра совмещён с уровнем распределения и называется уровнем свёрнутого ядра.

Уровень доступа

Уровень доступа, как правило, представляет собой локальную сеть, состоящую из хостов сети и коммутатора второго уровня, также могут быть использованы беспроводные сетевые устройства. Также на данном уровне активно используются услуги IP-телефонии.

Уровень распределения

На данном уровне обычно устанавливают коммутаторы 3-го уровня, которые осуществляют маршрутизацию между VLAN'ами, а также изолируют от потенциальных проблем в сети между зданиями кампуса, сохраняя сеть. Коммутаторы или маршрутизаторы на данном уровне являются центральными в топологии, связывая друг с другом WAN's и LAN's. Поэтому важно производить на данном уровне агрегирование каналов и создавать избыточность для максимальной минимизации ущерба от сбоев на уровне.

Уровень ядра

Главная задача данного уровня – связать здание с интернетом, другими зданиями и WAN'ами. Такая магистраль должна обеспечивать высокую скорость коммутации и маршрутизации пакетов. Аналогично уровню распределения, необходимо предусмотреть избыточность каналов связи и устройств для обеспечения отказоустойчивости, масштабирования, а также такого сервиса как QoS. Поэтому аппаратура должна использоваться высокопроизводительная, позволяющая работать на гигабитных скоростях.

Заключение

В последние 10-15 лет кампусные сети развиваются активными темпами. Каждый год создаются новые аппаратные и программные решения. Для реализации принципов построения сети используются иерархическая структура сети. Одним из основоположников данной концепции является международная телекоммуникационная компания Cisco. На базе их структуры SONA (Cisco Service-Oriented Network Architecture) в настоящее время строятся большинство кампусных сетей.

Библиографический список

1. Таненбаум Э. Компьютерные сети. 5-е изд. — СПб.: Питер, 2012. — 960с.
2. Олифер В., Олифер Н. Компьютерные сети. Принципы, технологии, протоколы. 5-е изд – СПб: Питер, 2016 – 992с.
3. Campus LAN and Wireless LAN DesignGuide – официальный сайт Cisco дата публикации: январь 2018 URL:<https://www.cisco.com/c/dam/en/us/td/docs/solutions/CVD/Campus/CVDCampus-LAN-WLAN-Design-Guide-2018JAN.pdf> (дата обращения 20.05.2018)
4. Голдовский Я.М. Проектирование кампусных сетей: Учебное пособие - М.: МИИТ, 2009 - 130 с

E.V. Denisov, A.V. Semashko

OPTIMIZING THE CONSTRUCTION OF CAMPUS NETWORKS

Nizhny Novgorod state technical university n. a. R. E. Alekseev

The paper describes the basic principles of building campus networks and provides a detailed analysis of the structure of campus networks. A three-level network model consisting of components is considered: the core level that provides communication with the "outside" world, i.e. access to the World Wide Web; the distribution level that provides routing within the campus between local networks and access to the organization's network services; the access level that provides access to the network on the campus.

Keywords: local area network, campus network, network architecture, access, resources, network structure.

УДК 621.313

Д.Ю. Кукишев, В.Н. Мещеряков, А.М. Евсеев, Д.С. Татаринев

РАЗРАБОТКА АЛГОРИТМА СОВМЕСТНОГО УПРАВЛЕНИЯ КОММУТАЦИЕЙ КЛЮЧЕЙ АВТОНОМНОГО ИНВЕРТОРА НАПРЯЖЕНИЯ И ИМПУЛЬСНЫМ ПРЕОБРАЗОВАТЕЛЕМ НАПРЯЖЕНИЯ В ЗВЕНЕ ПОСТОЯННОГО ТОКА

Липецкий государственный технический университет

Наиболее часто используемым методом формирования токов в фазах в современных преобразователях частоты является широтно-импульсная модуляция (ШИМ). Алгоритм разработан для условия постоянства выпрямленного напряжения на входе автономного инвертора напряжения (АИН), при изменении этого напряжения алгоритм будет более сложен в реализации и потребует добавочных вычислительных мощностей. В связи с этим актуальными являются разработки электропривода с упрощенными алгоритмами управления, такими, как релейное управление транзисторными ключами АИН, с обеспечением регулирования напряжения, подаваемого на вход инвертора в преобразователе частоты с помощью дополнительного коммутатора в звене постоянного тока, что позволяет улучшить качество напряжения питания статора двигателя при сохранении стандартной частоты коммутации ключевых элементов АИН.

Ключевые слова: инвертор напряжения, асинхронный двигатель, звено постоянного тока, поддержание напряжения.

В данной статье рассмотрим систему управления асинхронным короткозамкнутым двигателем (АДКЗ) с использованием транзисторного ключа в цепи постоянного тока инвертора напряжения. Введение транзисторного ключа позволяет регулировать напряжение, подаваемое на вход инвертора. В качестве управляемого транзисторного коммутатора будет использоваться IGBT-транзистор. Силовая схема импульсного понижающего преобразователя напряжения приведена на рисунке 1.

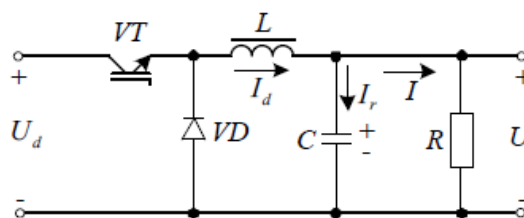


Рис. 1. Силовая схема понижающего преобразователя напряжения

Когда коммутирующий транзистор VT включен, конденсатор C и дроссель L накапливают энергию от источника питания. При выключении транзистора энергия, запасенная в дросселе L, отдается в нагрузку и конденсатор C.

Управление коммутацией IGBT-транзистора происходит релейным регулятором. В случае использования понижающего преобразователя в качестве стабилизатора выходного напряжения должен быть реализован скользящий режим управления выходному напряжению. Оно поддерживается постоянным около некоего заданного значения. Функциональная схема блока управления коммутатором (БУК) показана на рисунке 2.

Здесь задание на среднее напряжение сравнивается с измеренным выходным значением напряжения U_{dc} , а затем через блок усиления подается на вход релейного регулятора. Частота включения транзисторного коммутатора при этом зависит от ширины гистерезиса блока релейного регулятора, а также заданным значением напряжения. Релейный регулятор выдает сигнал управления (gtp) коммутатором VT. Благодаря этому уровень напряжения стабилизируется около заданного.

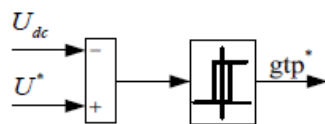


Рис. 2. Блок управления коммутатором

В ходе работы выделяются два коммутационных промежутка. На первом транзистор VT включен, источник постоянного напряжения подключен к нагрузке через дроссель L. При этом дроссель L и конденсатор C накапливают энергию от источника, которая также расходуется на нагрузке R. Схема замещения для данного режима представлена на рисунке 3.

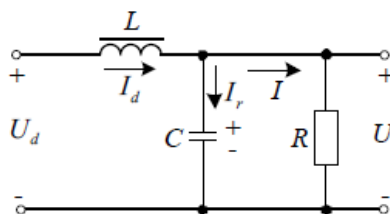


Рис. 3. Схема замещения понижающего преобразователя при включенном коммутационном транзисторе

На втором коммутационном промежутке транзистор VT выключается и ток с дросселя L замыкается через диод, показанный на рисунке 3.18, а также ток конденсатора протекают на выход преобразователя напряжения в нагрузку R. Схема замещения при выключенном коммутационном транзисторе показана на рисунке 4.

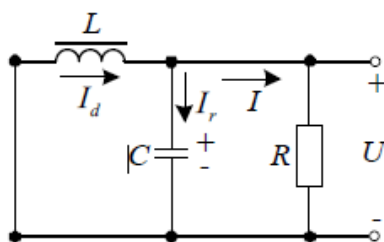


Рис. 4. Схема замещения понижающего преобразователя при выключенном коммутационном транзисторе

Функциональная схема системы управления асинхронным электроприводом со встроенным БУК в понижающем преобразователе напряжения показана на рисунке 3.22. Она построена на базе функциональной схемы системы управления, показанной на рисунке 3.6, но все блоки вычислений и преобразователь для упрощения не показаны, а есть лишь блок векторного управления, который получает обратные связи по току, скорости и положению ротора и выдает управляющие воздействия на ключи инвертора напряжения.

Силовая схема асинхронного электропривода со встроенным понижающим преобразователем показана на рисунке 5. Преобразователь (2) встраивается в звено постоянного тока после неуправляемого выпрямителя (1), а затем инвертор напряжения (5).

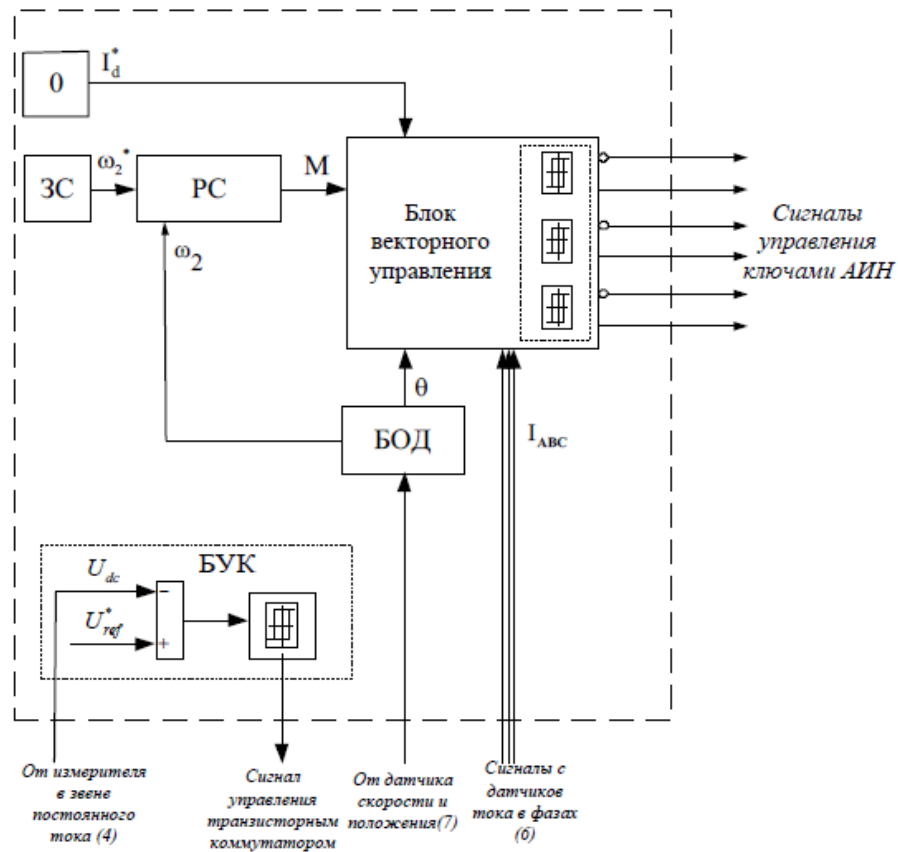


Рис. 5. Функциональная схема системы векторного управления асинхронным электроприводом с понижающим преобразователем в звене постоянного тока

БУК содержит замкнутый контур с узлом суммирования, который рассчитывает задание на релейно-гистерезисный регулятор напряжения, учитывая задание на напряжение и измеренное значение напряжения. Регулятор выдает управляющий сигнал на включение/отключение транзисторного коммутатора. Таким образом, реализуется скользящий режим по напряжению. В случае падения уровня сигнала на входе регулятора ниже заданного порогового, транзисторный коммутатор VT включается, при превышении порогового уровня выше заданного транзисторный коммутатор отключается.

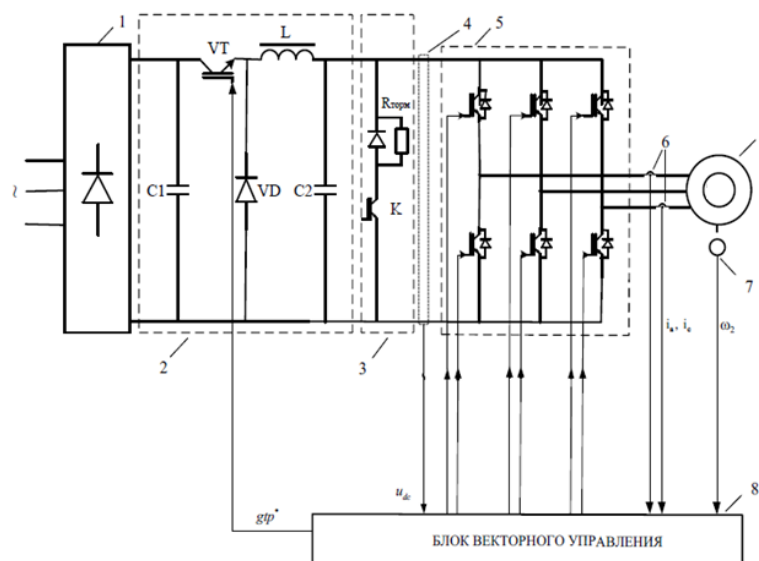


Рис. 6. Силовая схема асинхронно электропривода с понижающим преобразователем напряжения

В соответствии с силовой и функциональными схемами построим модель асинхронного электропривода в среде Matlab Simulink. Проведя моделирование, можно отметить, что после коммутатора в звене постоянного тока подключен импульсный понижающий преобразователь напряжения, работающий в режиме стабилизации.

На рисунке 6 показана модель асинхронного электропривода с понижающим преобразователем в среде Matlab при управлении с поддержанием разных уровней задания на напряжение в различные промежутки времени:

- $t_1=0 \dots 0,1c$ $U_{dc}^*=350B$;
- $t_2=0,1 \dots 0,21c$ $U_{dc}^*=400B$;
- $t_3=0,21 \dots 0,4c$ $U_{dc}^*=450B$;
- $t_4=0,4 \dots 0,8c$ $U_{dc}^*=500B$.

Обратная связь на релейный регулятор приходит от измерителя напряжение в звене постоянного тока. Ширина модуля гистерезиса для релейного регулятора в звене постоянного тока выбрана равной 1В. Результаты моделирования приведены на рисунках 7-9.

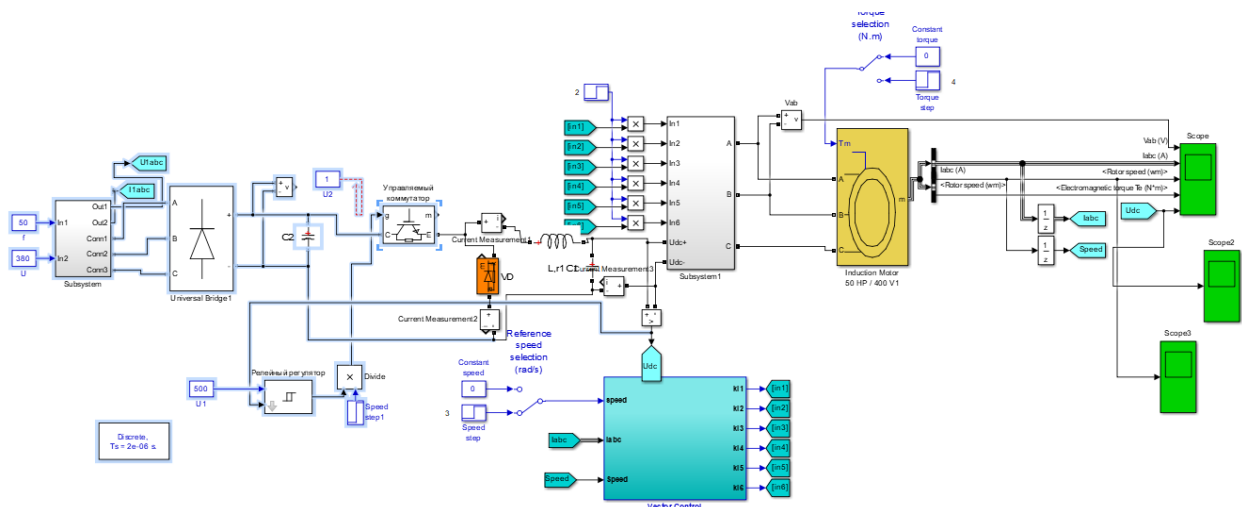


Рис. 7. Модель асинхронного электропривода с понижающим преобразователем в звене постоянного тока

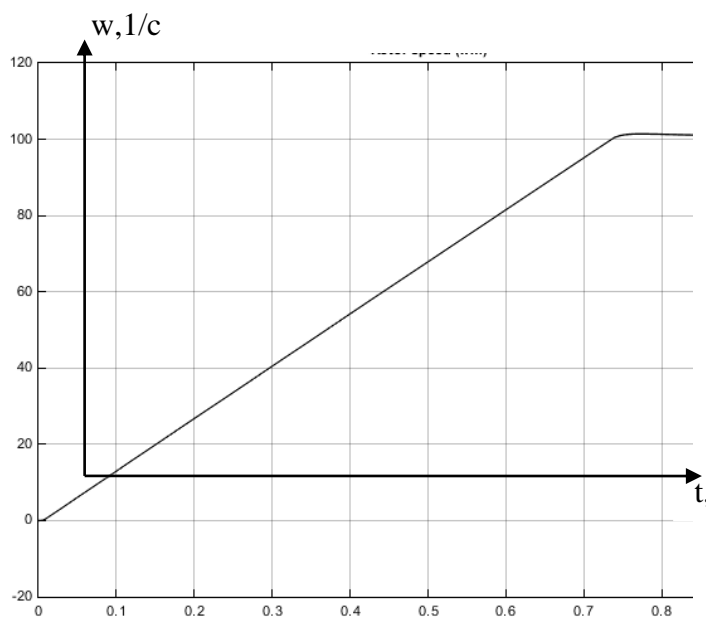


Рис. 8. График изменения скорости

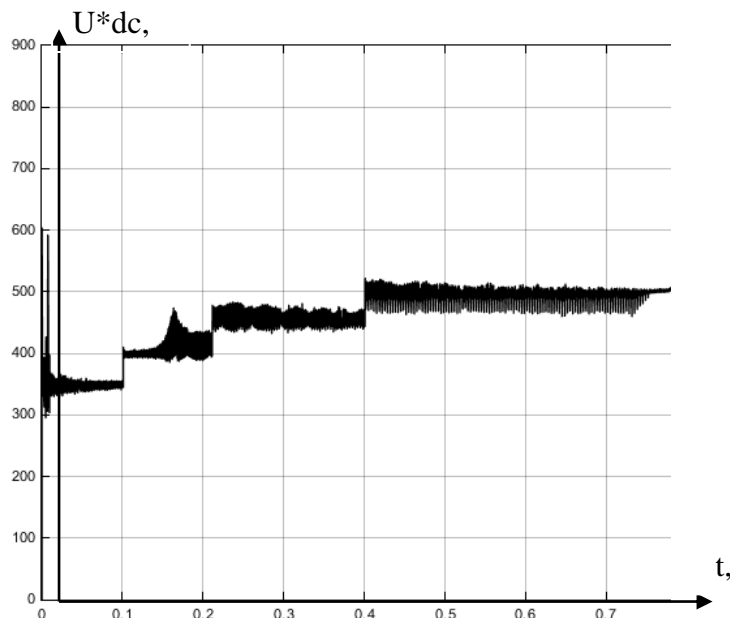


Рис. 9. График напряжения в звене постоянного тока

В данном случае понижающий преобразователь представляет собой не что иное, как стабилизатор напряжения с хорошими свойствами. Поэтому необходимо выбирать данный метод управления, как наиболее рациональный, в случае требования применений стабилизатора напряжения. В случае применения асинхронного электропривода в нефтяной промышленности, требуется повышение и поддержание напряжения на выходе преобразователя, поэтому проведем исследование импульсного повышающего преобразователя напряжения.

Статья публикуется при поддержке гранта РФФИ 19-48-480001 «Разработка, исследование и оптимизация энергосберегающих электротехнических и электроприводных автоматизированных комплексов для плазменных, электрометаллошлаковых и индукционных технологий и агрегатов».

Библиографический список

1. Усынин Ю.С. Системы управления электроприводов: Учеб. Пособие. – Челябинск: Изд-во ЮУрГУ, 2001. - 358 с.
2. Патент №166655. Устройство для управления электроприводом переменного тока// Мещеряков В.Н., Воеков В.Н. / Патентообладатель федеральное государственное бюджетное образовательное учреждение высшего профессионального образования «Липецкий государственный технический университет» (ЛГТУ) (RU) - №2016113199/07; заявл. 06.04.2016; опубл. 10.12.2016. Бюл. №34.
3. Соколовский Г.Г. Электроприводы переменного тока с частотным регулированием / М.: Academia. 2006. – 259 с.
4. Мещеряков В.Н. Оптимизация взаимного положения векторов тока статора и магнитного потока асинхронного двигателя при векторном управлении / В.Н. Мещеряков, П.Н. Левин // Известия вузов. Электромеханика. – 2006. – №1. – С.25-27.
5. Мещеряков В.Н., Абросимов А.С. Анализ систем управления электроприводом на базе автономного инвертора тока с релейными регуляторами тока и релейными регуляторами напряжения со сглаживающим емкостным фильтром. – Электротехнические комплексы и системы управления 2011. №3. С.64-68.
6. Мещеряков В.Н., Ласточкин Д.В. Анализ инверторов тока и напряжения с системами релейного управления / Вести вузов Черноземья. 2016. №2 с 13 -18.

7. Черных И. В. Моделирование электротехнических устройств в MATLAB, SimPowerSystems и Simulink / И. В. Черных. – М.: ДМК Пресс; СПб.: Питер, 2008. – 288 с.
8. Барышников О. Д. Построение систем автоматизированного электропривода / О. Д. Барышников, Г. Г. Соколовский, В. А. Новиков, В. М. Шестаков. – Л.: ЛДНТП. – 1968.
9. Башарин А. В. Управление электроприводами / А.В. Башарин, В.А. Новиков, Г. Г. Соколовский. – Л.: Энергоиздат. – Ленингр. отд-е. – 1982. – 392.

D.Yu. Kukishev, V.N. Meshcheryakov, A.M. Evseev, D.S. Tatarinov

**DEVELOPMENT OF AN ALGORITHM FOR JOINT CONTROL
OF SWITCHING OF KEYS OF AN AUTONOMOUS VOLTAGE INVERTER
AND A PULSE VOLTAGE CONVERTER IN THE DC LINE**

Lipetsk State Technical University

The most commonly used method for generating phase currents in modern frequency converters is pulse width modulation (PWM). The algorithm was developed for the condition of the constancy of the rectified voltage at the input of the autonomous voltage inverter (AVI), when this voltage changes, the algorithm will be more difficult to implement and will require additional computing power. In this regard, the development of an electric drive with simplified control algorithms, such as relay control of transistor switches of the AVI, with the provision of regulation of the voltage supplied to the input of the inverter in the frequency converter using an additional switch in the DC link is relevant, which improves the quality of the voltage power supply of the motor stator while maintaining the standard switching frequency of the AVI key elements.

Keywords: voltage inverter, asynchronous motor, DC link, voltage maintenance.

В.Н. Мещеряков, В.В. Пикалов, А.М. Евсеев, О.В. Данилова, Д.В. Ласточкин

ЭЛЕКТРОДУГОВОЙ ПЛАЗМОТРОН С ЭЛЕКТРОПРИВОДНОЙ СИСТЕМОЙ ДВИЖЕНИЯ ИСПОЛНИТЕЛЬНЫХ МЕХАНИЗМОВ

Липецкий государственный технический университет

Анализ известных разработок электродуговых плазмотронов показал, что наиболее перспективными с точки зрения простоты и работоспособности представляются конструкции, в которых зажигается электрическая дуга, поддерживается ее непрерывное горение и в зону горения подается плазмообразующий газ. Выполненные эксперименты показали ее работоспособность. На основании анализа литературных источников и собственного опыта экспериментальных исследований осуществлен выбор материалов для электродов

Ключевые слова: электрометаллургия, плазмотрон, электрическая дуга, электропривод, система управления, синхронный двигатель с постоянными магнитами.

В опубликованных работах в области электроплазменных установок [1-3] рассмотрены происходящие в плазме процессы, оптимальные температурные режимы, выбор материала электродов, необходимых для создания электрической дуги, состав плазмообразующего газа и ряд других вопросов, связанных с процессом горения плазмы. Однако в основном рассмотрены только опытные установки и результаты проведенных исследований при практически ручном управлении объектом без надлежащей автоматизации всех процессов.

Известные принципы и системы автоматизации электрометаллургических процессов могут быть применены при создании автоматизированной системы управления электродуговыми плазмотронами различной конструкции. Система управления питанием электродов, например, в сварочном производстве, обычно выполняется с замкнутым контуром управления током с ПИ-регулятором. В системе управления блоком поддержания заданного расстояния между электродами в условиях их постепенного сгорания целесообразно применить наиболее экономичные сервоприводы, построенные на базе вентильного синхронного двигателя с постоянными магнитами и датчиком положения ротора с системой частотного управления. Наиболее экономичные системы управления индуктором целесообразно выполнять на базе активного выпрямителя. Частотно регулируемым электроприводом, рекомендуемым для большинства систем вентиляции. Для повышения электромагнитной совместимости электрооборудования плазмотрона с питающей сетью целесообразно применять активный фильтр электроэнергии.

Задачей разработки предложенной конструкции электродугового плазмотрона является обеспечение регулирования потока плазмы в пространстве для обработки плоских поверхностей деталей, что может быть достигнуто за счет воздействия магнитного поля и регулирования скорости движения, температуры и объема плазмы на выходе трубчатого корпуса.

Разработанная конструкция электродугового плазмотрона содержит трубчатый корпус 1, выполненный из непроводящего ток тугоплавкого материала, имеющего в стенке сквозные расположенные соосно перпендикулярно оси корпуса отверстия, в которых находятся анодный электрод 2 и катодный электрод 3, подключенные к выходам блока питания 4 с регулируемым по уровню и постоянным по знаку напряжением. Один торцевой конец трубчатого корпуса соединен с узлом подачи рабочего плазмообразующего газа 5. В стенке трубчатого корпуса выполнены каналы для

прохождения охлаждающего агента 6. Соосно с трубчатым корпусом установлен подвижный трубчатый магнитопровод 7, внутренние размеры полости которого больше наружных размеров трубчатого корпуса 1. В полости подвижного трубчатого магнитопровода 7 между его внутренней поверхностью и внешней поверхностью корпуса соосно расположены два полюса 8 и 9 с обмотками 10, выводы которых подключены к источнику регулируемого напряжения постоянного тока 11, причем ось полюсов 8 и 9 расположена перпендикулярно по отношению к оси положения электродов 2 и 3. К выводам анодного электрода 2 и катодного электрода 3 также подключены выводы высоковольтного блока зажигания дуги 12.

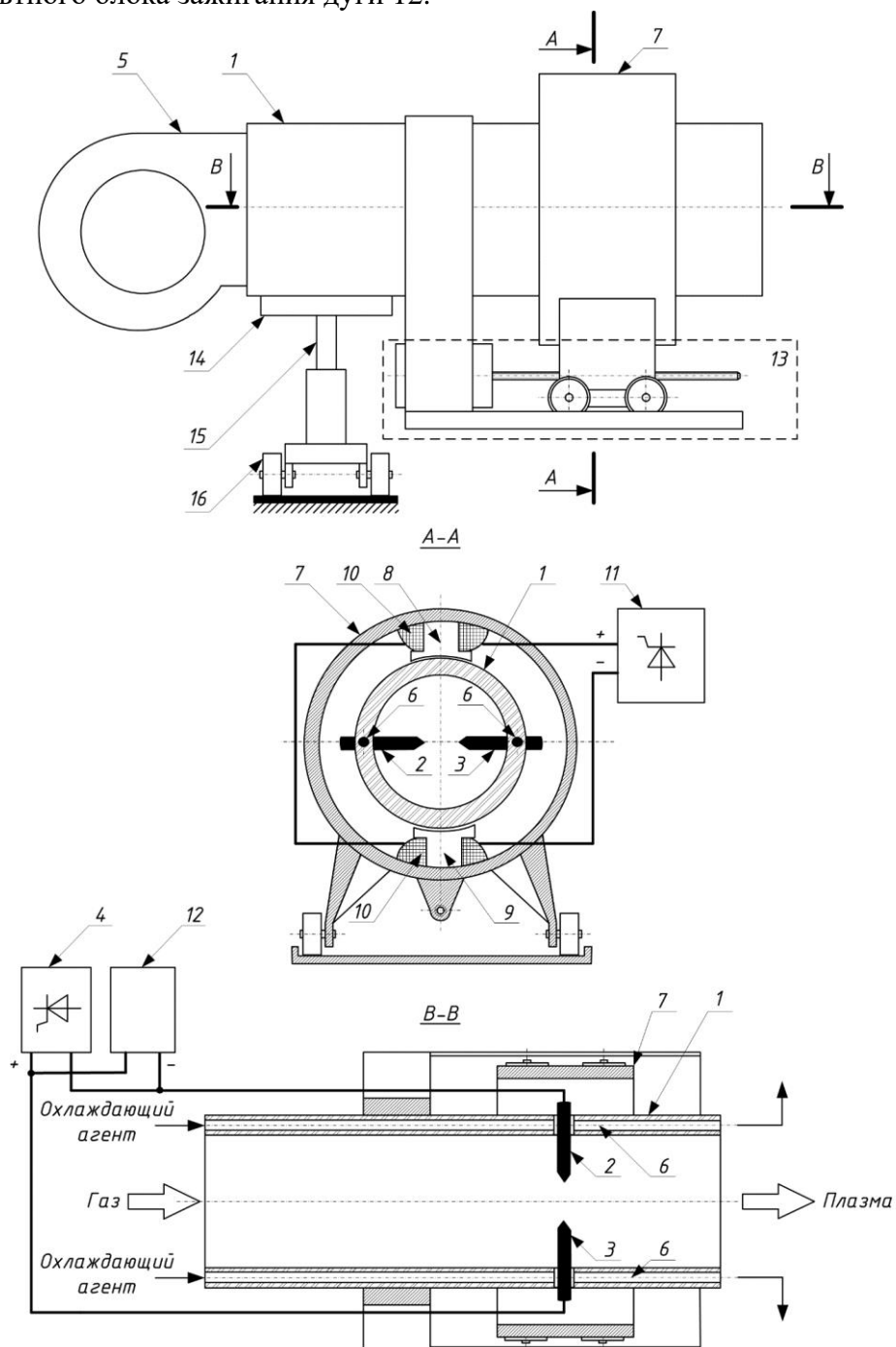


Рис. 1. Конструкция электродугового плазмотрона

Подвижный трубчатый магнитопровод 7 с полюсами 8 и 9 с обмотками 10 соединен с приводом 13, обеспечивающим его возвратно-поступательное движение вдоль

корпуса 1. Трубчатый корпус 1 закреплен на подвижной платформе 14, имеющей приводы возвратно-поступательного движения вверх-вниз 15, вправо-влево 16.

Электродуговой плазмотрон работает следующим образом.

К аноду 2 и катоду 3 от блока питания 4 подводится напряжение и иницируется электродуговой разряд и зажигается дуга. В полость трубчатого корпуса 1 к его торцевому входу от узла подачи 5 подается плазмообразующий газ под давлением, который проходит через дугу, и ионизируется с образованием плазмы, которая выходит из второго торцевого вывода трубчатого корпуса. Под действием движущегося плазмообразующего газа дуга растягивается. Для предотвращения ее разрыва подключают обмотку возбуждения 10, установленную на полюсах 8 и 9, к выходам источнику регулируемого напряжения постоянного тока 11, ток возбуждения протекает по обмотке 10 и создает электромагнитное поле, возникает электромагнитная сила, действующая на дугу, в направлении, противоположном направлению движения плазмообразующего газа, и стабилизирующая положение дуги.

При необходимости перемещения дуги вдоль продольной линии внутри цилиндрического корпуса 1 включают привод 13, обеспечивающий возвратно-поступательное движение вдоль корпуса 1 трубчатого магнитопровода 7 с полюсами 8 и 9 и с обмотками 10.

При увеличении скорости движения и расхода плазмообразующего газа увеличивают напряжение на выходе источника 11, соответственно, увеличивают ток возбуждения в обмотке 10, вследствие чего увеличивается величина электромагнитного потока, создаваемого полюсами 8 и 9. В результате возрастает электромагнитная сила, действующая на дугу в зоне ее горения в направлении, противоположном направлению движения плазмообразующего газа, стабилизирующая положению дуги и препятствующая ее разрыву. Увеличение расхода плазмообразующего газа позволяет увеличить количество вырабатываемой плазмы. При необходимости увеличения мощности, выделяемой в дуге, увеличивают напряжение, подводимое от блока питания 4 к выводам анода 2 и катода 3, при этом возрастает ток, протекающий через дугу, возрастает температура и результирующая мощность выработанной плазмы. При необходимости перемещения дуги вдоль продольной линии внутри цилиндрического корпуса 1 включают привод 13, обеспечивающий возвратно-поступательное движение вдоль корпуса 1 трубчатого магнитопровода 7 с полюсами 8 и 9 и с обмотками 10. Для перемещения зоны действия плазмы перемещают подвижную платформу 14 с закрепленным на ней трубчатым корпусом 1 вверх-вниз или вправо-влево, с помощью регулируемых электроприводов возвратно-поступательного движения 15 и 16. При этом обеспечивается обработка плоской поверхности плазмой.

Основными переменными плазмотрона являются температура, напряжение дуги, ток дуги, интенсивность теплового излучения. Соотношения и зависимости между этими переменными зависят в первую очередь от условий горения разряда в плазмотроне, которые рассмотрены в работе [4].

Проверка работоспособности предложенной конструкции плазмотрона была выполнена на экспериментальной установке, приведенной на рисунке 2.

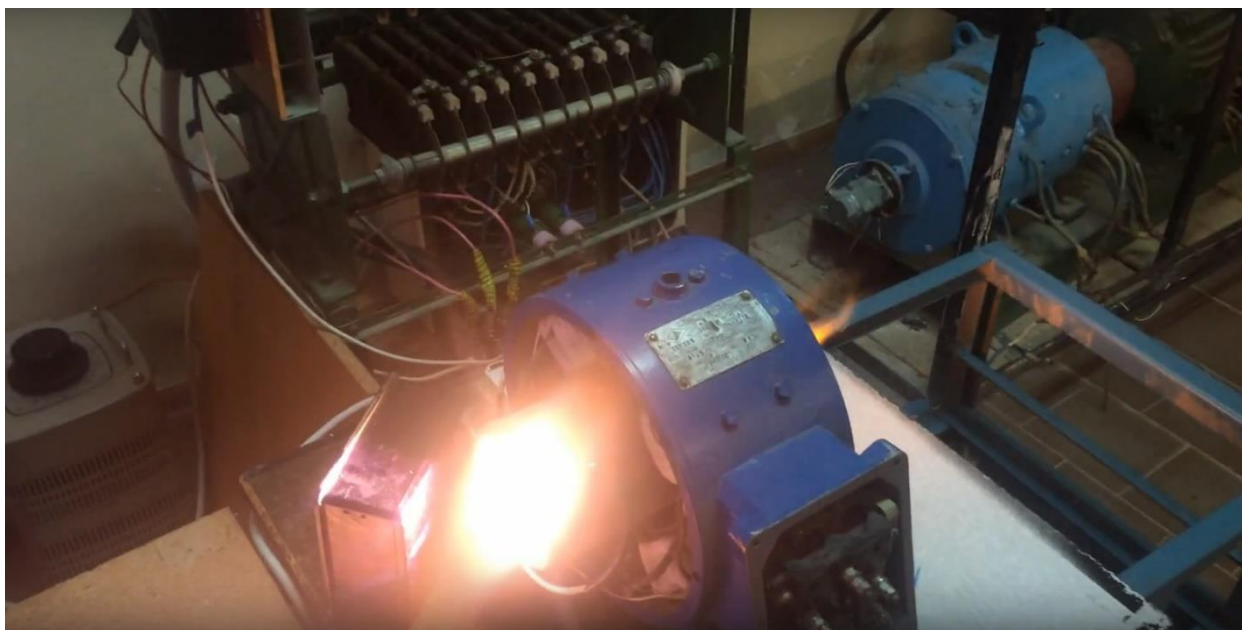


Рис. 2. Общий вид экспериментальной установки

Одной из основных особенностей разработанной конструкции плазмотронов является расположение электродов друг против друга, что упрощает поджиг дуги, стабилизацию ее положения и возможность вытягивания внешним электромагнитным полем. Оба электрода выполнены из графита круглого сечения с диаметром 10 мм.

Уравнение баланса напряжения в силовой цепи с электрической дугой постоянного тока при условии приведения всех параметров к звену постоянного тока имеет вид [5]

$$R_{\Sigma} I + L_{\Sigma} \frac{dI}{dt} = E - \alpha - \beta(l + \Delta l), \quad (1)$$

где I - ток дуги; R_{Σ} - активное сопротивление силовой цепи; L_{Σ} - индуктивность силовой цепи; E – ЭДС силовой цепи, приведенная к звену постоянного тока; α - приэлектродное падение потенциалов; β - градиент потенциала в столбе дуги; l - длина дуги; Δl - изменение длины дуги.

В электродуговых установках необходимо обеспечивать стабильное горение электрической дуги между электродами, питающимися от источника постоянного напряжения, что достигается за счет стабилизации тока дуги

Управление процессом горения дуги осуществляется путем регулирования напряжения питания обмотки электромагнита, определяющего величину протекающего через нее тока.

Уравнение баланса напряжений в обмотке электромагнита в переходном режиме имеет вид

$$U_{\text{э}} = I_{\text{э}} R_{\text{э}} + L_{\text{э}} \frac{dI_{\text{э}}}{dt}, \quad (2)$$

где $U_{\text{э}}$ - напряжение питания обмотки электромагнита; $I_{\text{э}}$ - ток в обмотке; $R_{\text{э}}$ - активной сопротивление обмотки; $L_{\text{э}}$ - индуктивность обмотки.

Постоянная времени обмотки электромагнита $T_{\text{э}} = L_{\text{э}} / R_{\text{э}}$, вследствие необходимости создания достаточно больших тянущих усилий магнитного поля, достаточно велика. При необходимости ускоренное изменение тока в обмотке может быть обеспечено за счет форсировки питающего напряжения. Магнитная цепь электромагнита должна быть ненасыщенной, в этом случае создаваемый магнитный поток прямо пропорционален току в обмотке.

В качестве исполнительного электропривода, обеспечивающего перемещение электродов 2 и 3 плазматрона по мере их сгорания, перемещение трубчатого магнитопровода 7, предлагается использовать электропривод на базе синхронного двигателя с постоянными магнитами (СДПМ).

Для эффективного управления СДПМ необходимо построить векторную систему подчиненного регулирования в осях (d, q), в контуре тока которой формируется составляющая тока статора по q-оси, создающей электромагнитный момент при постоянном магнитном потоке ротора. Ток по d-оси поддерживается равной нулю, поскольку не создает электромагнитного момента, но может вызвать потери в двигателе, а также позволяет правильно ориентировать вектор тока статора, чтобы поддерживать угол 90 градусов между ним и ротором [6,7]. Функциональная схема системы векторного управления ЭП с совместным управлением транзисторными ключами АИН и импульсным преобразователем напряжения представлена на рис.6.3.

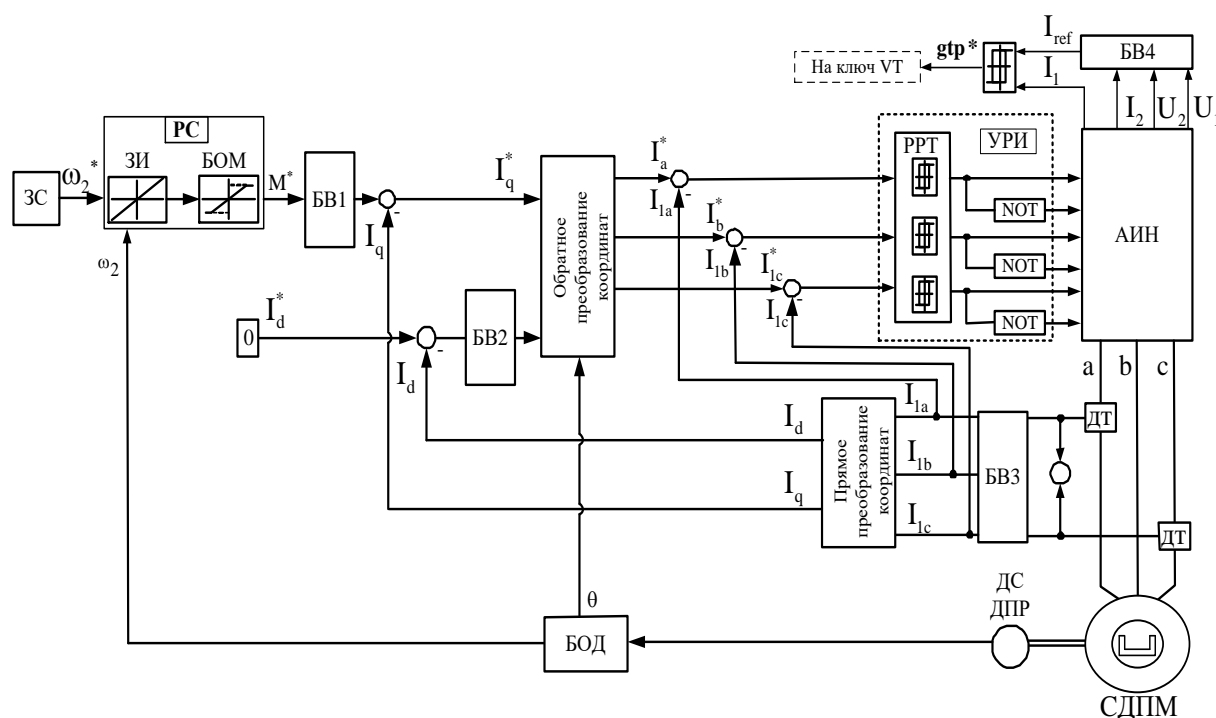


Рис. 3. Функциональная схема векторной системы управления электроприводом

Представлены следующие функциональные блоки: ЗС – задатчик скорости; РС – регулятор скорости, в котором содержатся блоки задатчика интенсивности (ЗИ) и блок ограничения момента (БОМ); БВД – производит оценку скорости и угла положения ротора двигателя; УПИ – устройство раздачи импульсов на силовые ключи АИН с релейными регуляторами тока; БВ1, БВ2, БВ3 – блоки вычислений, в которых производятся расчеты. Блоки прямого и обратного преобразования координат необходимы для перехода из системы (А, В, С) к системе (d, q) и обратно. В блоке вычислений БВ4 реализовано выражение для поддержания баланса мощности и формируется сигнал сравнения I_{ref} .

Рассмотренная конструкция электродугового плазматрона характеризуется простотой, обеспечивает регулирование скорости движения и расхода плазмы на выходе трубчатого корпуса, а также перемещение зоны действия плазмы.

Работа выполнена при поддержке гранта РФФИ 19-48-480001 «Разработка, исследование и оптимизация энергосберегающих электротехнических и электроприводных автоматизированных комплексов для плазменных, электрометаллошлаковых и индукционных технологий и агрегатов».

Библиографический список

1. А.с. СССР № 599732. Электродуговой нагреватель газа постоянного тока / Жуков М.Ф., Лыткин А.Я., Худяков Г.Н., Аньшаков А.С. Опубл. 07.09.1982. Бюл. № 33.
2. Патент РФ № 2465748. Электродуговой плазматрон / Мчедалов С.Г. Опубл. 27.10.2012. Бюл. № 30.
3. Патент на полезную модель № 188618. Электродуговой плазматрон / Мещеряков В.Н., Евсеев А.М., Пикалов В.В., Чупров В.Б., Конев В.А. Опубл. 18.04.2019. Бюл. №11.
4. Патент на изобретение RU 2715054 С1. Электродуговой плазматрон Мещеряков В.Н., Конев В.А., Евсеев А.М., Пикалов В.В., Конев М.В., 25.02.2020.
5. Жуков М.Ф., Смоляков В.Я. Урюков Б.А. Электродуговые нагреватели газа (плазматроны). М.: Наука. 1973. - 232 с.
6. Доманов В.И. Влияние случайных факторов на работу системы регулирования тока дуговой плавильной установки / В.И. Доманов, А.В. Доманов, К.Е. Карпухин и др. // Приводная техника. - 2010. №6, С. 16-19.
7. Воеков В.Н. Векторная система управления вентиляльным электроприводом на базе автономного инвертора напряжения с релейным регулированием входного тока инвертора и фазных токов статора. / Воеков В.Н., Мещеряков В.Н. // Вестник ЮУрГУ. Серия «Энергетика». – 2017. – Т.17 - №2. - С. 48-57.

V.N. Meshcheryakov, V.V. Pikalov, A.M. Evseev, O.V. Danilova, D.V. Lastochkin

ELECTRIC ARC PLASMA TORCH WITH ELECTRIC DRIVE SYSTEM OF MOTION OF EXECUTIVE MECHANISMS

Lipetsk State Technical University

The analysis of the known developments of electric arc plasmatrons showed that the most promising from the point of view of simplicity and operability are structures in which an electric arc is ignited, its continuous combustion is maintained, and a plasma-forming gas is supplied to the combustion zone. The experiments carried out have shown its efficiency. Based on the analysis of literature sources and our own experience of experimental research, the choice of materials for the electrodes was carried out.

Keywords: electrometallurgy, plasmatron, electric arc, electric drive, control system, permanent magnet synchronous motor.

СИСТЕМА СНИЖЕНИЯ ПУСКОВЫХ ТОКОВ И МОМЕНТА ОДНОФАЗНОГО АСИНХРОННОГО ДВИГАТЕЛЯ

Липецкий государственный технический университет

Подключение к сети однофазного конденсаторного двигателя сопровождается возникновением пиковых бросков пускового тока и слабозатухающих колебаний электромагнитного момента. Данные переходные процессы характерны для прямого пуска всех типов асинхронных двигателей. Эти броски и колебания приводят к износу изоляции обмоток, биению, резонансным колебаниям механического оборудования. Поскольку однофазные конденсаторные двигатели чаще всего применяются в системах с повторно-кратковременным режимом работы и без сложной оценивающей электроники (бытовые электроприборы, маломощные насосные станции), то они оказываются подвержены описанным негативным влияниям практически постоянно. Возникает задача создания простого дешевого устройства, позволяющего осуществлять запуск конденсаторных двигателей со сниженными пусковыми токами и малыми колебаниями электромагнитного момента. В статье рассматривается возможный алгоритм работы подобного устройства, основанный на задержке подачи напряжения на вспомогательную обмотку двигателя. Приведены результаты моделирования в *Matlab Simulink*.

Ключевые слова: конденсаторный двигатель, однофазный двигатель, прямой пуск, компьютерная модель, колебания момента, пусковой ток.

Введение

Значительная часть маломощных механизмов (насосные станции, бетономешалки, деревообрабатывающие станки) подключается напрямую к бытовой однофазной сети и приводится в движение посредством однофазных асинхронных двигателей. Стоит отметить, что, несмотря на то, что работа двигателя переменного тока возможна даже от одной подключенной фазы, для его запуска необходимо как минимум две обмотки [1]. С этой целью в однофазных двигателях применяется вторая (пуская) обмотка, которая при выходе на установившуюся скорость чаще всего отключается. Таким образом во время разгона однофазный двигатель является по сути двухфазным [2]. Как известно, для устойчивой работы у двухфазного двигателя обмотки смещены относительно друг друга на 90° , между фазными токами также должен поддерживаться сдвиг фаз на 90 электрических градусов. Для его создания в однофазных двигателях пусковая обмотка соединяется параллельно с основной через фазосмещающий конденсатор, потому другое название однофазных асинхронных двигателей – конденсаторные [3].

Двигатели такого типа устанавливаются на механизмы лишь на основании мощностных параметров и минимальной цены всех компонентов. Для избегания удорожания электроприводы выполняют нерегулируемыми, подключают к сети прямым пуском и выполняют лишь минимальные функции защиты от перегрузки и короткого замыкания. При этом во время запуска переходные процессы токов фаз характеризуются большими пиковыми значениями, а переходный процесс электромагнитного момента представляет собой слабозатухающие колебания [1, 4]. Эти недостатки приводят к просадке напряжения, ускоренному износу изоляции обмоток, а в сочетании с большой инерционностью привода – к ударам, биению и резонансным колебаниям механического оборудования [5-6]. Потому снижение электромеханических нагрузок при прямом пуске является актуальной задачей для конденсаторных двигателей. Возможным способом решения этой проблемы является создание простого неуправляемого устройства, осуществляющего коммутацию пусковой обмотки в строго определенный момент времени, в зависимости от мгновенного значения синусоидального тока первой обмотки [7-8].

Модель однофазного двигателя

Исследование, направленное на разработку подобного устройства, проводилось путем моделирования в Matlab Simulink. В качестве объекта исследования был взят блок Single Phase Asynchronous Machine из библиотеки SimPowerSystem, работающий в режиме раздельного управления основной и вспомогательной обмотками. Данная модель безусловно имеет допущения, так обмотки статора и ротора распределены равномерно и создают идеально синусоидальные намагничивающие силы, не учитываются зубцовые пульсации потока, потери от вихревых токов учтены в виде сопротивлений потерь в стали [9]. Кроме того, реальный двигатель, соответствующий этой модели, выпускается в США и имеет нетипичную для нас размерность, например, номинальную частоту 60 Гц и номинальное напряжение 0,25 л.с. (лошадиных сил). Полный перечень параметров модели двигателя приведен в табл. 1.

Графики электромагнитного момента, угловой скорости и модуля двухфазного тока при прямом пуске двухфазного двигателя показаны на рис. 1.

Таблица 1

Параметры модели однофазного двигателя

Параметр	Значение
Номинальная мощность, л.с. (Вт)	0,25 (187)
Номинальное напряжение статора, В	110
Номинальная частота, Гц	60
Сопротивление первой обмотки статора, Ом	2,02
Индуктивность первой обмотки статора, мГн	7,4
Сопротивление второй обмотки статора, Ом	2,02
Индуктивность второй обмотки статора, мГн	7,4
Сопротивление обмотки ротора, Ом	4,12
Индуктивность обмотки ротора, мГн	5,6
Взаимная индуктивность обмоток, мГн	177,2
Момент инерции, кг·м ²	0,0146
Коэффициент трения	0
Число пар полюсов	2
Коэффициент соотношения обмоток статора	1

Способ ограничения пусковых токов и момента был ранее описан для трехфазного [7] и двухфазного двигателя [8]. Однако, в данных статьях исследовалась сама возможность подобного воздействия на привод. Задержка подачи напряжения на одну из фаз в компьютерных моделях осуществлялась в самих блоках источников переменного тока как время, равное четверти периода синусоиды. В данном же исследовании мы разработали достаточно простой алгоритм, способный осуществлять подобную задержку на практике.

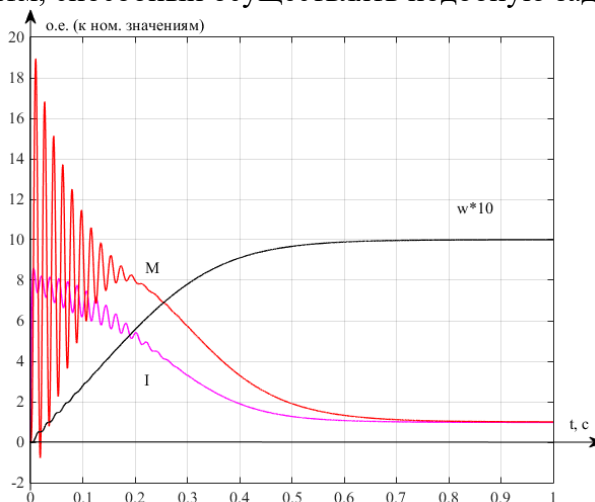


Рис. 1. Переходные процессы прямого пуска однофазного двигателя

Результаты моделирования алгоритма работы системы задержки подачи напряжения

Пиковые броски возникают при наложении нескольких мгновенных значений токов, подаваемых одновременно. Упомянутый способ ограничения пусковых токов основан на том, что сначала следует подавать напряжение на одну фазу, а когда его мгновенное значение будет минимальным (нулевым) – подключать вторую фазу. При этом порядок чередования двух фаз определяет направление вращения вала двигателя. Для направления «вперед» вторая фаза должна опережать первую.

Разработанная структурная схема устройства задержки напряжения показана на рис. 2.

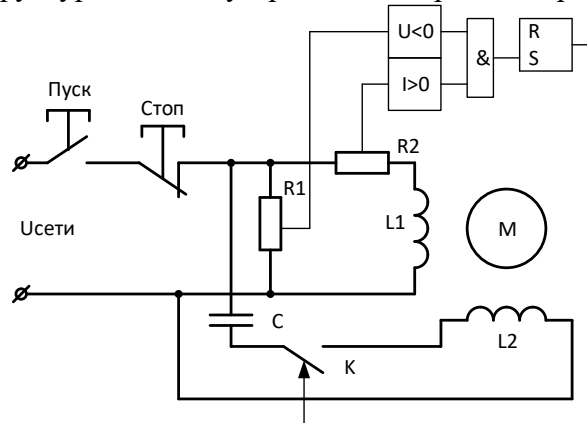


Рис. 2. Структурная схема системы задержки подачи напряжения

Как отмечалось выше, момент коммутации пусковой обмотки должен происходить в момент приближения мгновенного значения напряжения основной обмотки к нулю, при этом движению «вперед» должна соответствовать положительная полуволна тока этой обмотки. Данные условия выполняются посредством сочетания работы двух компараторов, сравнивающих значения входных напряжений с нулем. В качестве сигнала синусоиды тока на компаратор подается напряжение, снятое с шунта, подключенного последовательно с первой обмоткой. Удержание выполнения данного логического условия обеспечивается *RS*-триггером, выходной сигнал с которого подается на коммутатор, в роли которого может выступать *MOSFET*-транзистор с высокой скоростью срабатывания. Таким образом, схемное решение устройства задержки подачи напряжения достаточно простое и включает в себя минимальное количество дополнительных элементов, которые могут быть размещены на небольшой плате с блоком питания.

На рис.3 показаны переходные процессы модуля тока, скорости и момента, а на рис. 4 – переходные процессы токов и напряжений, полученные при моделировании однофазного двигателя с задержкой подачи напряжения на пусковую обмотку статора.

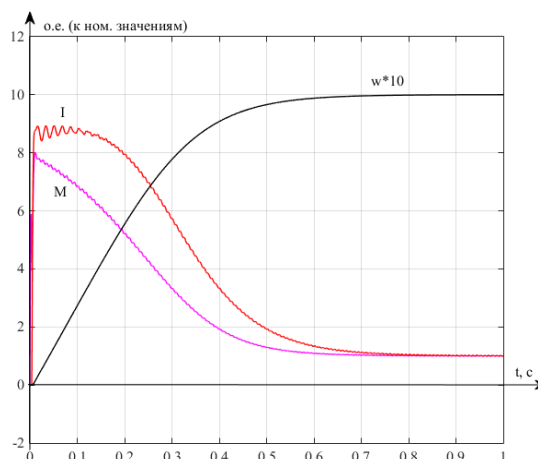


Рис. 3. Переходные процессы однофазного двигателя с разработанной схемой подключения

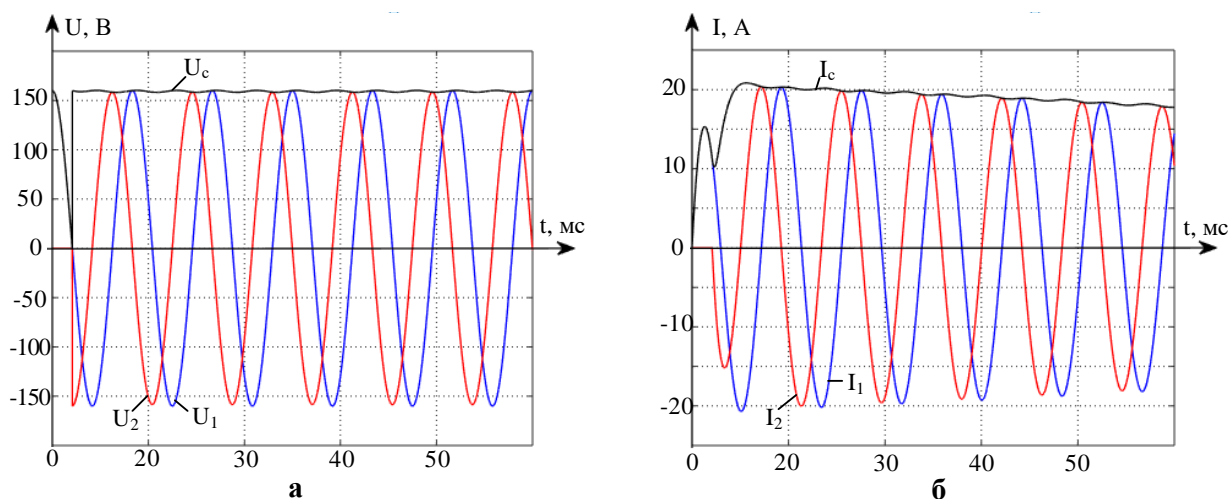


Рис. 4. Переходные процессы напряжения (а) и токов (б) однофазного двигателя с задержкой

Как видим на графиках, по сравнению с прямым пуском время разгона двигателя не изменилось и по-прежнему составляет 700 мс, а максимальные мгновенные значения токов и момента уменьшились за счет значительного сглаживания колебаний во время переходного процесса. Схожесть графиков тока и напряжения с описанными в предыдущей статье, где задержка формировалась как время задержки самого источника напряжения, говорит нам о правильности работы алгоритма разрабатываемой системы снижения колебаний. Компараторы точно определяют наиболее оптимальную точку коммутации, а триггер не позволяет электронному контактору размыкаться сразу после прохождения этой точки.

Заключение

Исследование подтвердило возможность ограничения пусковых токов и колебаний момента однофазного двигателя. Разработанная система задержки коммутации подтвердила работоспособность алгоритма, основанного на определении момента смены знака мгновенного значения тока основной фазы. Система, обеспечивающая выполнение данного алгоритма, может быть выполнена на основе простых электронных компонентов, питаемых от небольшого нерегулируемого блока питания. Для этого понадобятся два компаратора. Один из них должен сравнивать с нулем сигнал напряжения основной обмотки двигателя, а другой – сигнал напряжения с шунта, подключенного в цепь этой обмотки. Формируемый системой сигнал коммутации должен подаваться на электронный ключ, основным условием выбора которого является быстродействие – не более 1 мс. Предполагается, что таким ключом может послужить полевой транзистор. Реальное создание такой системы ограничения и эксперименты с ее подключением к однофазному двигателю будет являться предметом дальнейших исследований.

Статья публикуется при поддержке гранта РФФИ 19-48-480001 «Разработка, исследование и оптимизация энергосберегающих электротехнических и электроприводных автоматизированных комплексов для плазменных, электрометаллошлаковых и индукционных технологий и агрегатов».

Библиографический список

1. Исмагилов Ф.Р. Диагностика конденсаторных однофазных асинхронных двигателей с учетом технологических и эксплуатационных факторов / Ф.Р. Исмагилов, И.Х. Хайруллин, О.А. Бойкова, Д.Ю. Пашали // Вестник Южно-Уральского государственного университета. Серия: Энергетика. – 2011. – № 34 (251). – С. 28-34.

2. Abdel-Rahim N., Shaltout, A. Operation of single-phase induction motor as two-phase motor / N. Abdel-Rahim, A. Shaltout // IEEE 28th Annual Conference of the Industrial Electronics Society. IECON 02. – 2002. – № 2. – С.967-972.
3. Иванов-Смоленский А.В. Электрические машины: Учебник для вузов / А.В. Иванов-Смоленский. – М.: Энергия, 1980. – 928 с.
4. Hayakwong E. Two-phase induction motor drive improvement for PV water pumping system / E. Hayakwong, V. Kinnares, C. Bunlaksanunusorn // 19th International Conference on Electrical Machines and Systems (ICEMS), IEEE. – 2016. – С. 1-6.
5. Abu-Rub H. High performance control of AC drives with MATLAB/Simulink models / H. Abu-Rub A. Iqbal, J. Guzinski – United Kingdom: John Wiley & Sons Ltd, 2012. – 482 с.
6. Мещеряков В.Н. Ограничение колебаний электромагнитного момента асинхронного двигателя при скалярном частотном управлении / В.Н. Мещеряков В.В. Данилов // Вестник Южно-Уральского государственного университета. Серия: Энергетика. – 2018. – Т. 18. – № 3. С. 88-97.
7. Омельченко Е.Я. Снижение пусковых токов и моментов асинхронного короткозамкнутого двигателя за счет последовательности фазных напряжений / Е.Я. Омельченко, А.Б. Лымарь, В.О. Танич, М.Ю. Петушков // Электротехнические системы и комплексы. – 2020. – № 2 (47). – С. 47-54.
8. Белоусов А.С. Снижение пусковых токов и моментов двухфазного асинхронного короткозамкнутого двигателя за счет задержек фазных напряжений / А.С. Белоусов, В.Н. Мещеряков // Электротехнические системы и комплексы. – 2020. – № 4 (49). – С. 42-48.
9. Collins E.R. Single-phase induction motor adjustable speed drive: Direct phase angle control of the auxiliary winding supply / E.R. Collins, H.B. Puttgen, W.E. Sayle // Conference Record of the 1988 IEEE Industry Applications Society Annual Meeting. – 1988 – С. 246-252.

V.N. Meshcheryakov, A.S. Belousov

SYSTEM FOR REDUCING STARTING CURRENTS AND TORQUE OF A SINGLE-PHASE INDUCTION MOTOR

Lipetsk State Technical University

Purpose: This article describes a research of the development of an algorithm for the operation of an electronic system for limiting starting currents and torque fluctuations of a single-phase induction motor, based on a delay in the supply of voltage to the starting winding.

Design/methodology/approach: The methodology consists in simulating of the direct start of a single-phase induction motor and starting the same motor with a delay in the switching of the second winding by means of a limiting system with a developed algorithm and comparing the results of this simulation with the results of the operation of a previously known model, where the delay was created by the voltage source.

Findings: The research showed the possibility of reducing the starting currents and reducing of the electromagnetic torque oscillations of a single-phase motor by connecting the second winding to the grid at a time determined by the limiting system. The operation algorithm of this system can be based on comparing the instantaneous voltage and current values of the main winding and determining the most appropriate moment in time.

Research limitations/implications: This research opens further prospects for the selection of the element base and circuit design for the implementation of the developed algorithm for the operation of the electronic system. After that, it will be necessary to test this system in series of experiments.

Originality/value: The developed algorithm for switching the starting winding using the limiting system showed good results, indistinguishable from starting with a voltage delay, carried out in time. This algorithm can be implemented on a simple element base that does not require the use of expensive sensors.

Keywords: capacitor motor, single-phase motor, direct start, computer model, torque fluctuations, starting current.

**СТЕНД ОТЛАДКИ И ПРОВЕРКИ ПОДСИСТЕМЫ ОБНАРУЖЕНИЯ
НЕШТАТНЫХ СОБЫТИЙ НА ЛИНЕЙНОЙ ЧАСТИ МНОГОИТОЧНОГО
МАГИСТРАЛЬНОГО ГАЗОПРОВОДА СРЕДСТВАМИ СИСТЕМЫ ЛИНЕЙНОЙ
ТЕЛЕМЕХАНИКИ**

Нижегородский государственный технический университет им. Р.Е. Алексеева

Рассматривается вопрос создания стенда, обеспечивающего в лабораторных условиях формирование данных, аналогичных измеряемым системой линейной телемеханики на действующем многоиточном магистральном газопроводе и моделирующего процессы в реальном масштабе времени с функцией задания аварийных событий.

Ключевые слова: телемеханика, моделирование нештатных ситуаций, транспорт газа, обработка информации.

При транспортировке газа по многоиточным магистральным газопроводам (ММГ) возникают нештатные события, не только затрудняющие транспортировку, но и способные нанести существенный экономический и экологический ущерб, а также привести к человеческим жертвам.

К особо тяжелым нештатным событиям (аварийным ситуациям) при транспортировке газа относятся:

- несанкционированная перестановка линейных кранов;
- несанкционированная перестановка перемычек между газопроводами с разными режимами транспорта;
- утечка газа или полный разрыв газопровода, с возможным взрывом и возгоранием.

Нештатная перестановка линейного крана или перемычки на начальном этапе не является разрушающим действием для ММГ, но способна нанести ущерб транспорту газа в части поставок потребителям. Однако, при определенном развитии событий, данные нештатные события способны привести к разрушению ММГ за счет превышения допустимого давления газа или за счет возникновения резкого перепада давления при открытии кранов.

Таким образом, раннее обнаружение нештатных событий на линейной части ММГ является одной из основных задач автоматизации процесса транспорта газа.

Для решения этих задач применяют системы автоматики на компрессорных станциях и системы телемеханики на линейной части ММГ [1], обеспечивающие дистанционный контроль параметров транспорта газа и управление газовыми потоками на контролируемом участке посредством дистанционного изменения его конфигурации. Поскольку аварийные события на магистральном газопроводе возникают редко, а последствия могут быть катастрофическими, то появляется необходимость в обучении и тренинге диспетчерского персонала для их распознавания и принятия срочных и адекватных мер по их локализации. Обучение проводится на имитаторах аппаратно-программных средств диспетчерского пункта, работающих в режиме реального времени [2]. Вторым, и наиболее оптимальным путем, является создание автоматизированных систем обнаружения нештатных событий, как автономных, так и функционирующих на базе имеющихся систем автоматики и телемеханики [3].

В 2017 году в филиале РФЯЦ-ВНИИЭФ «НИИИС им. Ю.Е. Седакова» была разработана подсистема обнаружения нештатных событий (ПОНС) на базе системы телемеханики УНК ТМ. Источником данных для ПОНС являются штатные средства телемеханики, не требующие применения специализированных датчиков. В ходе

разработки, одним из наиболее сложных, стал вопрос об отладке и лабораторных испытаниях данной подсистемы, поскольку проведение натуральных экспериментов на действующем ММГ сопряжено с существенными организационно-техническими ограничениями. Решением данного вопроса стало создание стенда, обеспечивающего формирование данных, аналогичных действующему ММГ, в реальном масштабе времени. Именно решению данной проблемы и посвящена настоящая работа.

Математическая модель стенда

В качестве основы программного обеспечения (ПО) для стенда отладки и отработки ПОНС было выбрано ПО математической модели (ММ) тренажера диспетчера линейного производственного управления (ЛПУ), разработанной в филиале РФЯЦ-ВНИИЭФ «НИИИС им. Ю.Е. Седакова». ПО тренажера диспетчера ЛПУ работает в операционной системе реального времени ОС QNX 4.25. Основными возможностями данной математической модели являются:

- воспроизведение и моделирование работы линейной части и компрессорных цехов в реальном масштабе времени;
- формирование и отработка следующих типов нештатных событий на линейной части и компрессорных цехах в любой комбинации:
 - несанкционированная перестановка кранов;
 - аварийный останов газоперекачивающего агрегата (ГПА) и пуск ГПА;
 - неполное закрытие или открытие любых кранов;
 - аварийные ситуации, возникающие при частичном или полном разрыве газопроводов.

В основу математической модели тренажера диспетчера был положен матричный метод моделирования, который обеспечил возможность моделирования ММГ в реальном масштабе времени.

В процессе разработки тренажера диспетчера ЛПУ особое место было уделено тестированию правильности работы. В 2006 – 2008 годах руководство диспетчерской службы ООО «Газпром трансгаз Югорск» проводило всестороннюю проверку метода моделирования по всем типам нештатных ситуаций, реализованных в тренажере диспетчера ЛПУ, формировала рекомендации по улучшению метода моделирования, представляла консультации и требуемые данные для тестирования. В результате максимальная погрешность моделирования реальных аварийных ситуаций не превышала 2%.

В дальнейшем ММ тренажера успешно применялась для лабораторной отладки систем автоматического управления (САУ) расходом газа на базе телемеханики УНК ТМ.

Доработка ПО тренажера диспетчера

В процессе создания стенда потребовалась доработка ПО ММ в части:

- реализации драйверов обмена данными с контроллером контролируемого пункта (КП) УНК ТМ по протоколу обмена Modbus/RTU;
- формирования шумовых изменений показаний датчиков давления газа;
- формирование перепадов давления при моделировании утечек газа.

Алгоритм передачи данных моделирования

Рассчитанные технологические характеристики моделируемого ММГ передаются в базу данных, где хранятся в виде параметров показаний датчиков давления, температуры газа и состояния кранов. При запросе от КП параметров его работы, драйвер Modbus/RTU считывает данные из базы данных, формирует ответное сообщение и передает его КП. Поскольку ОС QNX позволяет использовать несколько драйверов параллельно, то ПО ММ было доработано для возможности функционирования с 8-ю драйверами протокола Modbus/RTU на скорости обмена 9600 бод для имитации информации 8-ми КП УНК ТМ. При этом драйвера ПО ММ работают в режиме Slave.

Структурная схема стенда предоставлена на рис. 1.

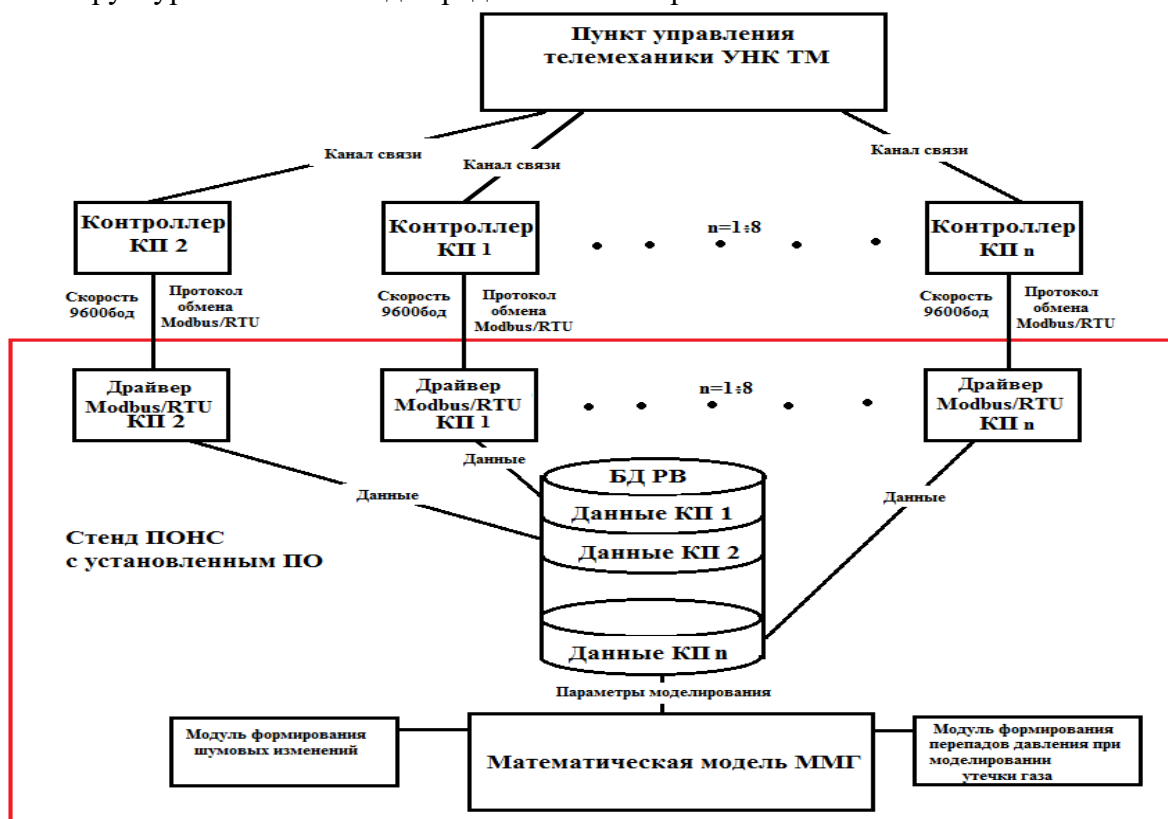


Рис. 1. Структурная схема стенда

Алгоритм формирования перепадов давления

Для формирования перепадов давления в ПО был использован алгоритм представления ММГ в виде весового неориентированного графа, где вершинами служат датчики давления, а весами – расстояния между этими датчиками. Дополнительной вершиной выступает место утечки/разрыва.

В ПО ММ был реализован алгоритм Дейкстры, основным назначением которого является поиск кратчайшего расстояния между вершинами неориентированного графа с весами с учетом состояния открытия кранов моделируемого ММГ.

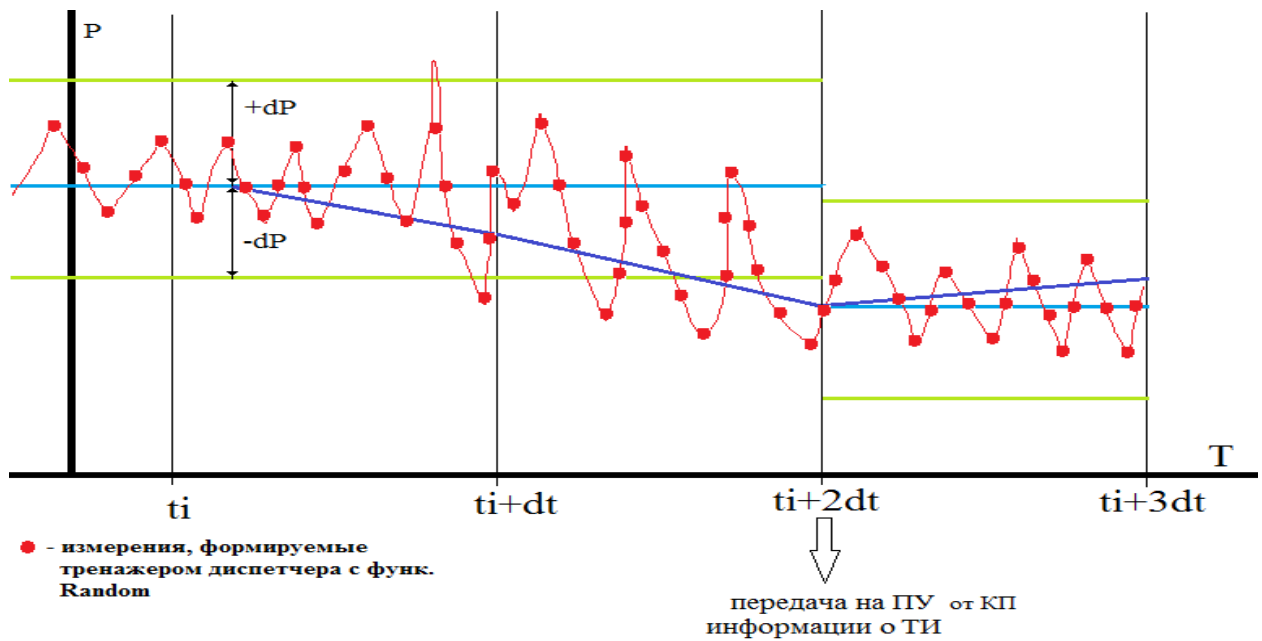
Учитывая, что скорость звука в метане около 430м/с, то время (Т), за которое перепад давления от места утечки/разрыва дойдет до датчика рассчитывается по формуле:

$$T = \frac{d}{1000 * 0.430} \quad (1)$$

Значения давления на датчиках остаются на уровне последнего вычисленного значения до начала моделирования разрыва, что обеспечивает моделирование прихода перепада давления для датчика.

Алгоритм формирования шумовых изменений

Для формирования шумовых изменений значений моделируемых датчиков для каждого датчика задается максимальный диапазон шумов по отношению к смоделированному значению и частота его возникновения, что обеспечивает более реальную картину для отработки подсистемы, представленную на рис. 2.

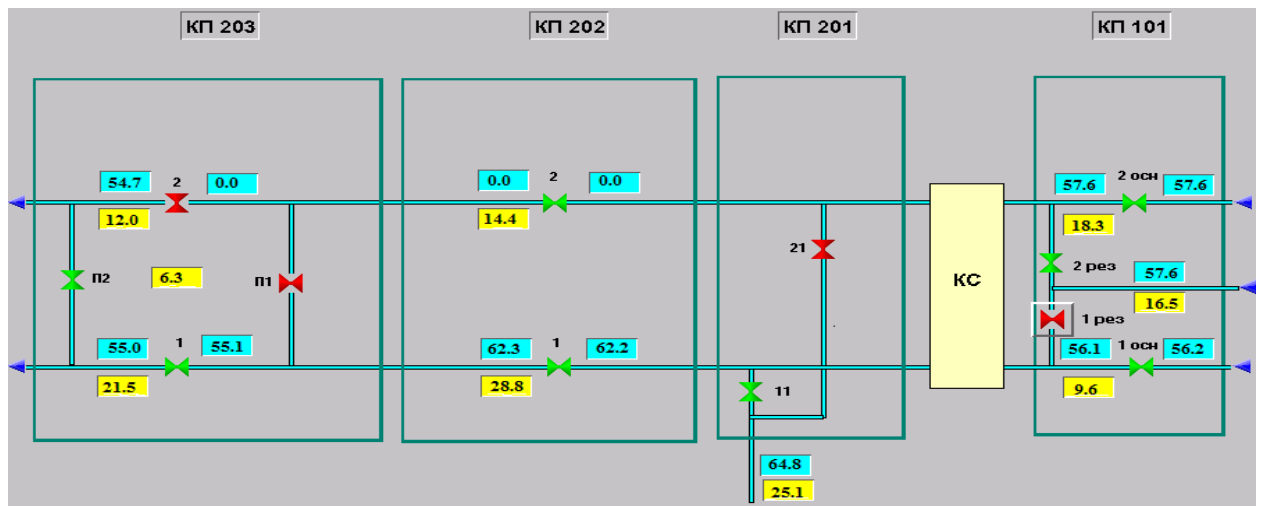


Испытание стенда

Конфигурирование математической модели

Для проведения испытаний стенда в соответствии с руководством на тренажер диспетчера ЛПУ была сконфигурирована математическая модель участка ММГ двухниточной системы ММГ, контролируемой системой телемеханики УНК ТМ, общей протяженностью моделируемого участка 230 км, включая компрессорную станцию.

Мнемосхемы участка моделирования приведены на рис. 3 – 4.



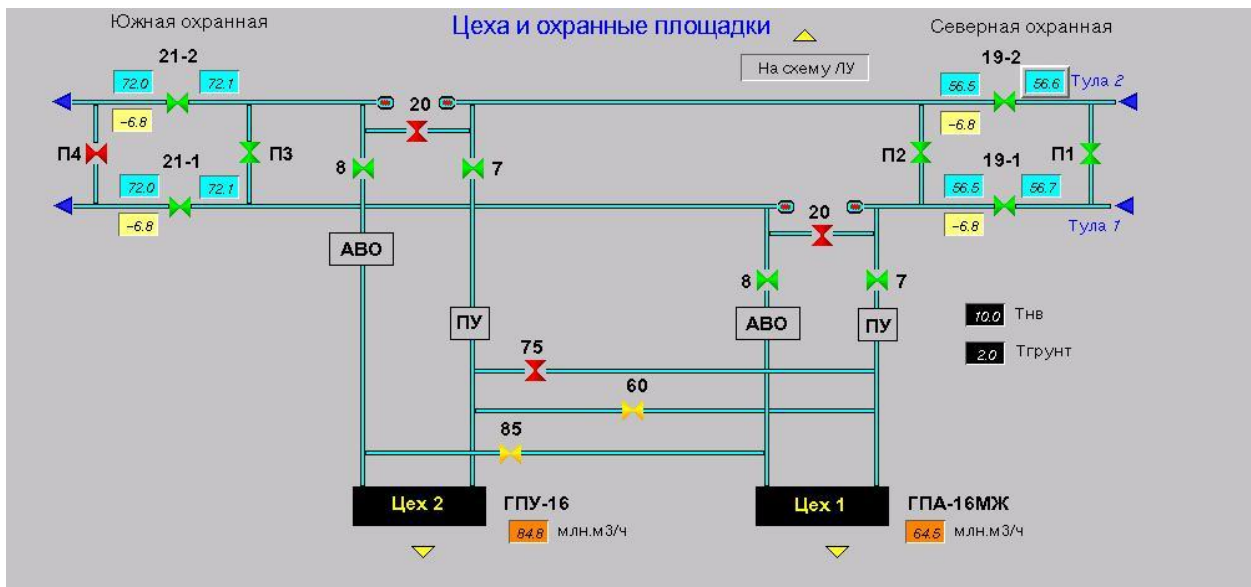


Рис. 4. Мнемосхема компрессорных цехов и охранных площадок

Для формирования перепадов давления была разработана карта (граф) моделируемого участка, представленная на рис. 5, и конфигурационный файл описания карты, состоящий из текстовых строк следующего формата: T1; T2; L; Кр; Key где: T1 и T2 – номера смежных вершин графа; L – расстояние между вершинами в метрах; Кр – номер КП в модели для крана, расположенного между вершинами; Key – номер крана в модели, расположенного между вершинами.

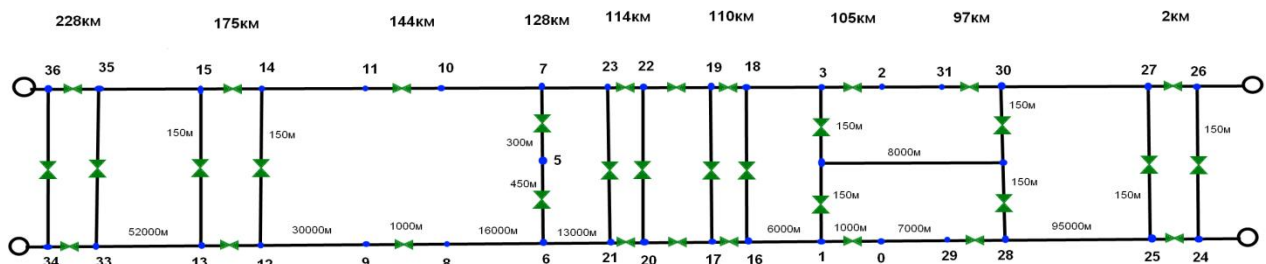


Рис. 5. Граф моделируемого ММГ с номерами вершин

Аппаратное обеспечение стенда

Для проведения испытаний стенда было подготовлено аппаратное обеспечение в составе:

- ПЭВМ с монитором, клавиатурой и манипулятором типа мышь;
- Коммуникационная плата BlueStorm/LP-4;
- 4 кабеля RS-232 типа 0-модем.

Модель и стенд были сконфигурированы для передачи информации на четыре КП телемеханики (рисунок 6).

В качестве имитаторов КП использовались контроллеры КП УНК ТМ с установленным ПО, работающем на реальном объекте. На ПЭВМ имитатора ПУ было установлено программное обеспечение пункта управления УНК ТМ работающее на реальном участке. Радиальная схема реального радиоканала системы телемеханики УНК ТМ имитировалась созданной сетью по каналу RS 485 [4].

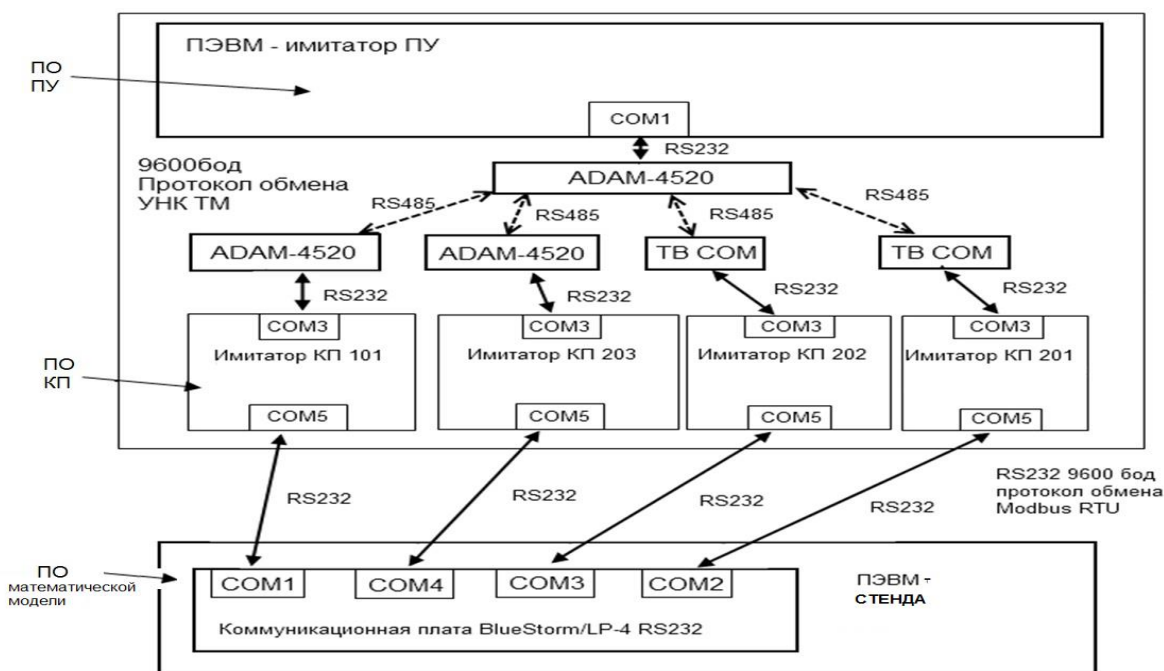


Рис. 6. Схема испытаний стенда

2.3. Моделирование перепадов давления

В модуле моделирования математической были сформированы несколько нештатных событий в виде утечки газа разной интенсивности в различных местах моделируемого ММГ. Данные сформированные математической моделью передавались на ПУ УНК ТМ через КП без искажений. На ПУ УНК ТМ фиксировались сформированные математической моделью стенда перепады давления с заданными характеристиками размеров. На рисунок 7 показаны тренды датчиков давления газа на участке моделирования.



Рис. 7. Графики датчиков давлений после разрыва газопровода

Для оценки адекватности моделирования проводилось сравнение трендов реальных утечек и трендов, сформированных математической моделью, при задании исходных данных, аналогичных параметрам реальной утечки. Сравнение показало, что погрешность моделирования датчиков давления при формировании аварии математической моделью не превышает 2% от реальных значений.

Выводы

Разработанный стенд с математической моделью реального времени, дополненный драйверами связи с КП по протоколу Modbus/RTU, модулями формирования шумов и перепадов давления газа, обеспечивает моделирование ММГ, нештатных событий на ММГ с точностью необходимой для отладки подсистемы обнаружения нештатных событий.

Библиографический список

1. Многоуровневые информационно-управляющие системы реального времени для топливно-энергетического комплекса России: Монография / Под ред. В.Е. Костюкова. – Нижний Новгород: Изд-во Нижегородского университета, 2007. – 243 с.
2. И.Р. Бухвалов. Методы и алгоритмы информационной поддержки управления газотранспортной системой: дис. канд. тех. наук: 05.13.06. – 2007.
3. С.В. Евсеев. Проблема определения теста утечки на многониточном магистральном газопроводе штатными средствами линейной телемеханики / ЭКСПОЗИЦИЯ НЕФТЬ ГАЗ Выпуск 6(73) ноябрь 2019. – С. 70 – 72.
4. И.Р. Бухвалов, С.В. Евсеев. Оптимизация передачи данных в системах телемеханики // Труды конференции «Информационные системы и технологии» ИСТ–2018. – С. 681 – 685.

I.R. Bukhvalov, S.V. Evseev, A.A. Zuev

OPTIMIZATION OF DATA TRANSFER IN TELEMECHANICS SYSTEMS

Nizhny Novgorod State Technical University n. a. R.E. Alekseev

The question of creating a stand that provides in laboratory conditions the formation of data similar to those measured by a linear telemechanics system on an operating multi-thread gas main pipeline and simulates processes in real time with the function of setting emergency events is considered.

Keywords: telemechanics (remote control), modeling of emergency situations, gas transport, information processing.

Д.А. Благодаров¹, Ю.М. Сафонов¹, О.В. Крюков²

НОВЫЕ КОНСТРУКТОРСКО-ТЕХНОЛОГИЧЕСКИЕ РЕШЕНИЯ В ЭНЕРГОСБЕРЕГАЮЩИХ ПРИВОДАХ

¹ Московский энергетический институт (ТУ), г. Москва,

² ООО «ТСН-электро», г. Нижний Новгород

Предложен анализ конструкций электроприводов с машинами переменного тока мегаваттного класса для интегрированных исполнений агрегатов нефтегазовой отрасли. Проведен обзор основных преимуществ капсулированных машин в электроприводах газоперекачивающих агрегатов. Рассмотрены особенности конструкций интегрированных электродвигателей с турбокомпрессорами на магистральных газопроводах России. Показаны эффективные результаты реализации безредукторных и безмасляных технологий для повышения функциональных возможностей, надежности, ресурса, энергоэффективности и экологичности наиболее ответственных электромеханических систем.

Ключевые слова: электродвигатель, конструкция, электрическая машина, ось вращения, надежность, электропривод, турбокомпрессор.

Введение

Перспективы дальнейшего повышения надежности и энергоэффективности основных технологических установок ТЭК связаны с реализацией всей совокупности инновационных достижений электроприводной техники в рамках мировых тенденций интеграции приводных электрических машин с исполнительными турбомеханизмами [1-3]:

- капсулированное исполнение,
- ликвидация редуктора (мультипликатора) и системы уплотнений,
- внедрение системы активных электромагнитных подшипников (ЭМП),
- использование устройств плавного запуска электроприводных агрегатов (ЭПА),
- подключение преобразователей частоты для регулирования скорости.

Следует учитывать, что проектирование и строительство новых электроприводных компрессорных станций (КС) сопровождается ужесточением условий работы агрегатов в более сложных климатических (северные регионы и шельф) и технологических ограничениях (морские платформы и подводные добычные комплексы). Это предполагает необходимость обеспечения новых возможностей по парированию стохастических возмущений [4-6], а также требований безопасной работы ЭПА в статике и динамике [7-9]. В этой связи уже сегодня согласно нормативной документации каждый ЭПА должен быть оснащен системами непрерывной диагностики [10-12] и прогнозирования технического состояния электропривода и агрегата с использованием технологий искусственных нейронных сетей [13-15].

Этим направлениям развития способствуют возможности современных электроприводных двигателей мегаваттного класса, которые связаны со сменой технологий электромашиностроения, силовой и микропроцессорной техники и переходом к интеллектуальным системам управления и мониторинга [16-18]. Однако наиболее существенные изменения произошли в конструкциях исполнительных машин и нагнетателей, что и позволило объединить их интегрированном (капсулированном) корпусе [19-21].

Обоснование целесообразности применения капсулированных ЭГПА

В современных ЭГПА именно характеристики приводной машины, ее энергетические и динамические характеристики, функциональные возможности, надежность и ресурс эксплуатации определяют конкурентоспособность и технико-экономические показатели транспорта газа по магистральным газопроводам (МГ) и его конкурентоспособность по сравнению с газотурбинными вариантами привода.

Отечественными предприятиями, связанными с разработкой высоковольтных машин для ЭГПА, являются ЗАО «РЭП Холдинг» (СПб) и ООО "Электротяжмаш-Привод" (г. Лысьва). Среди зарубежных фирм мегаваттные двигатели производят АВВ, консорциум Alstom (Converteam) – MAN Turbo, MELCO (MitsubishiElectricCorporation), GeneralElectric (GE), Siemens. Опыт промышленной эксплуатации данных машин в нефтегазовом комплексе для мощных и энергоёмких механизмов доказал перспективность их применения.

Разработка новых конструкций приводных электродвигателей ЭГПА обусловлена новыми требованиями надежности, энергоэффективности и экологичности на КС МГ в соответствии с отраслевыми нормативными документами. Модернизация и замена конструкций приводных двигателей, существующих ЭГПА на новые интегрированные с исполнительным турбомеханизмом, позволяет получить конкурентные преимущества:

- Исключение из компоновки установок ЭГПА мультипликатора.
- Прямое соединение «двигатель-нагнетатель» на базе высокоскоростных электрических машин.
- Значительное сокращение площадей под установку ЭГПА.
- Исключение системы смазки подшипников двигателя и компрессора и необходимости запасов масла, его охлаждения и рисков возгорания.
- Применение ЭМП роторов повышает показатели надежности и ресурс ЭГПА, снижаются затраты на обслуживание и ремонты.
- Снижение износа механического и электротехнического оборудования благодаря снятию ограничений на число пусков и остановов.
- Снижение вероятности возникновения поломок и помпажа, благодаря плавному изменению режимов работы компрессорных цехов.
- Повышение готовности агрегата к пуску, снижение времени и энергетических затрат пуска.
- Повышение точности отработки задания при дистанционном управлении ЭГПА с переходом к безлюдным технологиям обслуживания оборудования и улучшение условий труда.

Анализ реализации капсулированных ЭГПА «МОПИКО»

Рассмотрим результаты работ по технико-экономическому обоснованию применения капсулированного ЭГПА «МОРИСО» (рис. 1) при реконструкции КС Карталинская ООО «Газпром трансгаз Екатеринбург».

В ходе предпроектной подготовки были выделены и проработаны следующие варианты реконструкции:

1. Замена существующих электродвигателей СТМ-4000 на комплектный электропривод КРЭП-6300 ХК «Привод» (двигатели СТД-6300, преобразователи частоты «Ансальдо-ВЭИ»).
2. Модернизация существующих электродвигателей (замена подшипников, изоляции и т.д.) с установкой регулируемого электропривода.
3. Установка трех (2+1) мотор-компрессоров «МОРИСО» вместо группы из пяти существующих СТМ-4000 (плавное регулирование производительности по одной нитке МГ).
4. Установка пяти (4+1) мотор-компрессоров «МОРИСО» вместо СТМ-4000.

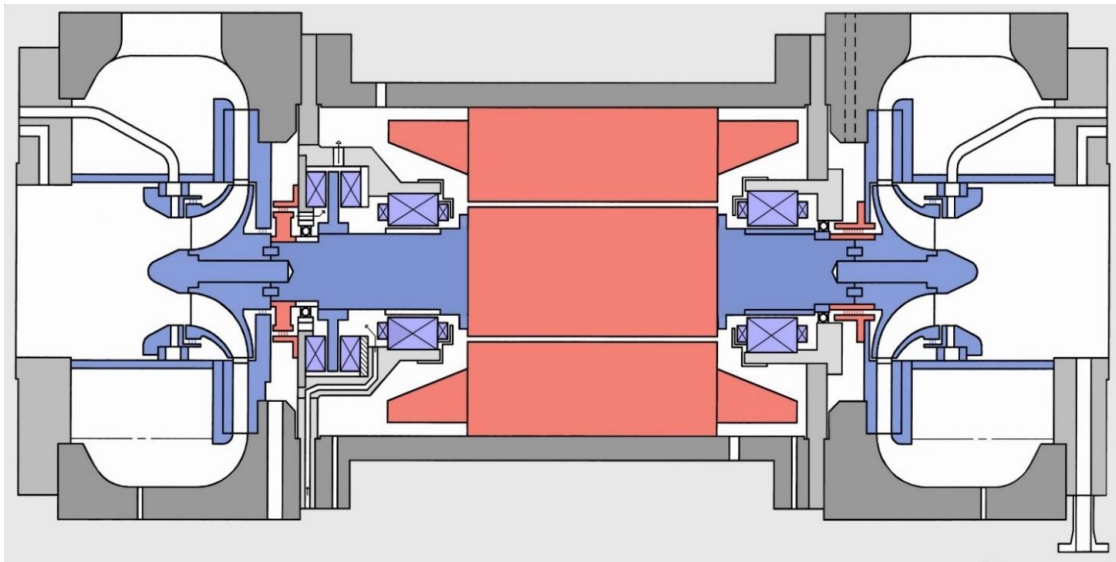


Рис. 1. Конструкция ЭГПА типа МОПИКО

Однако наиболее перспективным признан вариант с установкой оригинальных капсулированных мотор-компрессоров «МОРИСО» мощностью 7500 кВт со скоростью вращения 12000 об/мин. Данный наиболее рациональный вариант реконструкции предполагает замену пяти существующих двигателей на три мотор-компрессора «МОРИСО» производства фирмы «Альстом» (Германия). ЭГПА «МОРИСО» состоит из АД и двух нагнетателей, выполненных в едином конструктиве на одном валу, с активным ЭМП ротора двигателя и валов нагнетателей. Охлаждение двигателя осуществляется перекачиваемым газом, позволяя:

- плавно регулировать производительность в широких пределах;
- минимизировать капитальные вложения при максимальных ТЭР;
- отказаться от систем маслосмазки и уплотнения в подшипниках;
- в два раза сократить площади под привод при реконструкции.

Результаты технико-экономического сопоставления вариантов ЭГПА по КС Карталинская МГ «Бухара-Урал» приведены на рис. 2.

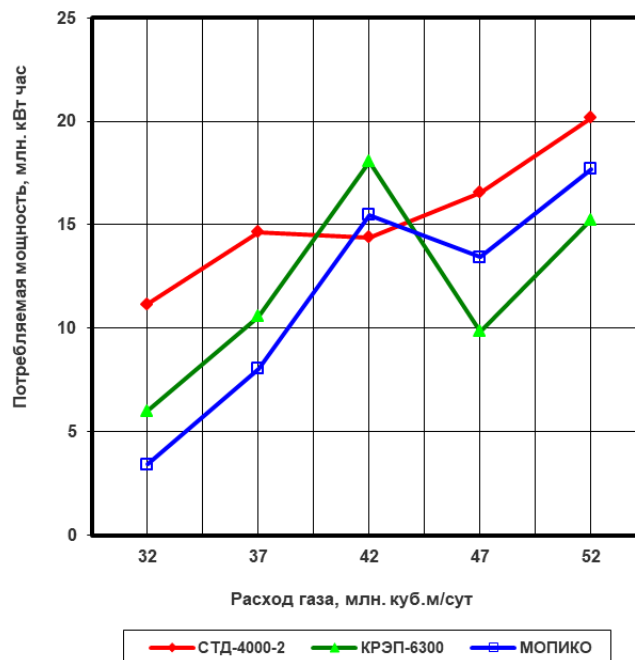


Рис. 2. Энергозатраты вариантов ЭГПА на КС Карталинская

Расчетные значения энергозатрат этих вариантов приведены в табл. 1. расчеты выполнены при условии работы гпа 6000 час/год там же получены итоговые значения энергозатрат для заданных суточных объемов газа в диапазоне значений от 32 до 52 млн. м³/сут (для каждого из расчетных значений принято равномерное распределение времени работы, принятое 1200 час).

Таблица 1

Годовые энергозатраты вариантов ГПА КС Карталинская

Оснащение ГПА	Годовые энергозатраты, млн. кВт·час	Годовой расход топливного газа, млн. м ³
СТД-4000-2	76,92	-
КРЭП-6300	59,70	-
ГПА-Ц-6,3	-	46,0
МОПИКО	58,08	-

Таким образом, применение капсулированных частотно-регулируемых ЭГПА более чем на 20 % уменьшает потребление электроэнергии КС по сравнению с нерегулируемыми системами, что значительно снижает нагрузку на питающие сети, стабилизирует выходное давление газа, снижает его себестоимость и надежность работы трубопроводной арматуры.

Система управления капсулированным ЭГПА

Эффективность применения ЭГПА с быстроходными асинхронными двигателями (АД) достигается в системах векторного управления с преобразователями частоты:

- на основе автономного инвертора тока и АД;
- на основе автономного инвертора напряжения и АД;
- по схеме «вентильный двигатель» с быстроходным СД.

Наиболее перспективным вариантом является частотно-регулируемый ЭГПА с автономным инвертором и быстроходным АД на магнитном подвесе. Схема управления капсулированного ЭГПА представлена на рис. 3.

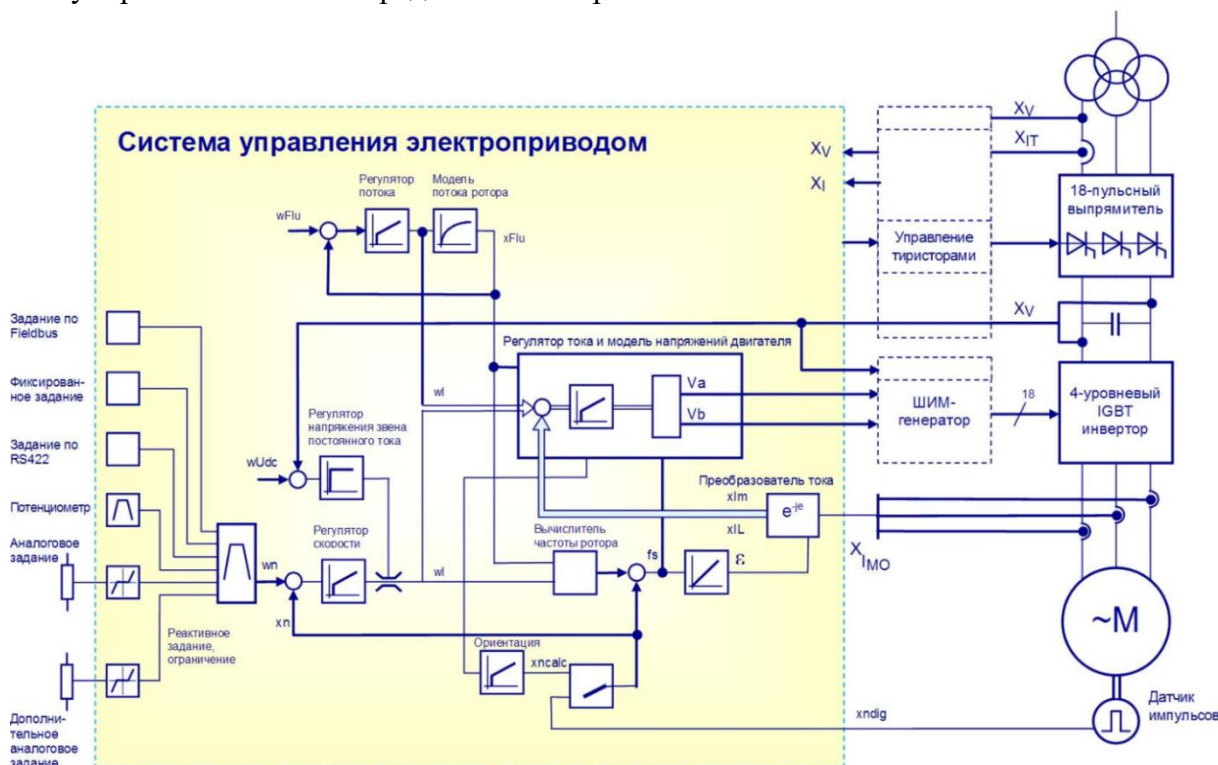


Рис. 3. Система управления капсулированным ЭГПА

Современный, высокоскоростной частотно-регулируемый ЭГПА «МОРИСО» выполнен на основе инвертора напряжения на интеллектуальных IGBT-модулях (1300 В, 1600 А), с оригинальным многоуровневым ШИМ-формированием напряжения АД. Применение новых силовых модулей совместно с конденсаторами, выполненных конструктивно с малыми индуктивностями рассеивания, позволяет ограничить пики напряжения на транзисторах.

Система управления "МОРИСО" выполнена по векторному принципу организации САР, ориентированной по вектору потокосцепления ротора. Очень важно, что современная мультипроцессорная САУ позволяет интегрировать привод в АСУ ТП и электроснабжением (АСУ ЭС) КС. Системные решения при реализации АСУ КС с ЭГПА позволяют оптимально и плавно регулировать производительность агрегатов с максимально возможными технико-экономическими показателями.

Интеграция электроприводов в АСУ ЭС через систему управления микропроцессорных терминалов релейной защиты и автоматики дает возможность автоматического дистанционного управления высоковольтными выключателями ЭГПА, позволяя:

- формировать оптимальные переходные процессы и снижать ударные значения моментов при резких и значительных изменениях нагрузки в МГ;
- обеспечить точный график газоподачи путём компенсации случайных возмущений климатического и технологического характера путём создания инвариантных систем с регрессионными алгоритмами;
- реализовать систему предстартовой и оперативной диагностики и прогнозирования неисправностей на базе нейро-сетей с эволюцией технического состояния ЭГПА по лингвистическим величинам текущих параметров;
- обеспечить функции стартовой самонастройки системы управления, а также возможностей переналадки и адаптации к новым режимам работы ЭГПА;
- получить экспертную поддержку для рационального выбора конкретного технического решения модернизации ЭГПА и комплексного исследования выходных характеристик на экспериментальном автоматизированном стенде.

Высокооборотный двухполюсный асинхронный электродвигатель с КЗ кованым высокопрочным ротором на ЭМП при номинальной скорости до 9.500 об/мин имеет отвод тепла посредством перекачиваемого газа, схема которого приведена на рис. 4.

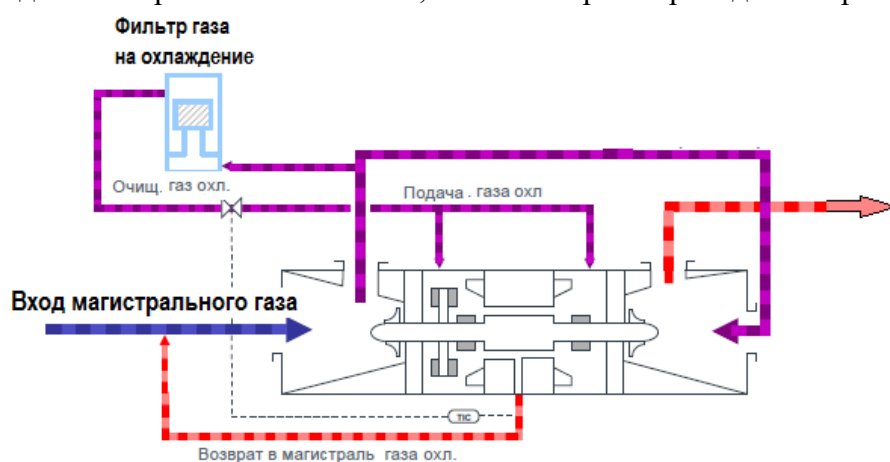


Рис. 4. Конструкция системы охлаждения моноблока ЭГПА

Выводы

Таким образом, новые конструктивные решения электродвигателей переменного тока мегаваттного класса в капсулированном исполнении позволяют значительно повысить показатели надежности, энергоэффективности и экологичности систем электроприводов агрегатов в важнейших электроэнергетических отраслях промышленности.

Библиографический список

1. Благодаров Д.А., Сафонов Ю.М., Крюков О.В. Энергосберегающие системы автоматизированного электропривода турбокомпрессоров // В сборнике: Информационные системы и технологии ИСТ-2020. XXVI МНТК. НГТУ им. Р.Е. Алексеева. 2020. С. 269-276.
2. Васенин А.Б., Крюков О.В. Проектирование электромеханической части и систем управления энергетических установок газотранспортных потребителей // Известия Тульского гос. ун-та. Технические науки. 2011. № 5-1. С. 47-51.
3. Туганов Р.Б., Благодаров Д.А., Крюков О.В. Перспективы автоматизации объектов машиностроения // В сборнике: Информационные системы и технологии ИСТ-2020. Сборник материалов XXVI МНТК. НГТУ им. Р.Е. Алексеева. 2020. С. 323-330.
4. Kryukov O.V., Blagodarov D.A., Dulnev N.N., Safonov Y.M., Fedortsov N.N. Intelligent control of electric machine drive systems // В сборнике: 2018 10th International Conference on Electrical Power Drive Systems, ICEPDS 2018 - Conference Proceedings 10. 2018. С. 8571670.
5. Мещеряков В.Н., Ласточкин Д.В., Крюков О.В. Приложения теории нечетких множеств для обработки данных и задач прогнозирования в системах АЭП // В сборнике: Современные сложные системы управления. Материалы XII МНПК. 2017. С. 153-158.
6. Крюков О.В. Алгоритмы быстрого преобразования Уолша в микропроцессорных системах управления электроприводом // Известия ВУЗов. Электромеханика. 2005. № 4. С. 39-44.
7. Васенин А.Б., Крюков О.В. Энергоэффективные и экологичные установки воздушного охлаждения // В сборнике: ВЕЛИКИЕ РЕКИ' 2017 труды научного конгресса 19-го МНПФ. НГАСУ. 2017. С. 93-96.
8. Крюков О.В., Степанов С.Е. Модернизация систем управления ЭГПА в условиях действующих компрессорных станций // В сборнике: Проблемы автоматизации и управления в технических системах. МНТК под редакцией М.А. Щербакова. 2013. С. 29-32.
9. Крюков О.В. Моделирование и микропроцессорная реализация электромеханических систем // Электротехника: сетевой электронный научный журнал. 2015. №3. С. 55-61.
10. Kryukov O.V., Gulyaev I.V., Terplukhov D.Y. Method for stabilizing the operation of synchronous machines using a virtual load sensor // Russian Electrical Engineering. 2019. № 7. С. 473-478.
11. Крюков О.В. Мониторинг условий эксплуатации электродвигателей газоперекачивающих агрегатов // Контроль. Диагностика. 2016. № 12. С. 50-58.
12. Крюков О.В., Мещеряков В.Н., Гуляев И.В. Электроприводы на основе машины двойного питания и АВК с преобразователями в цепях статора и ротора // Саранск, 2020.
13. Степанов С.Е., Васенин А.Б., Крюков О.В. Цифровое моделирование синхронных электрических машин при работе на турбокомпрессорную нагрузку // Наука и техника в газовой промышленности. 2020. № 4 (84). С. 84-90.
14. Васенин А.Б., Степанов С.Е., Крюков О.В. Методология и средства оперативного мониторинга электродвигателей на КС // Контроль. Диагностика. 2019. № 11. С. 52-58.
15. Благодаров Д.А., Крюков О.В. Повышение надежности энергетических установок путем прогнозирования состояния электроприводов // В сборнике: Великие реки' 2018 Труды научного конгресса. Отв. редактор А.А. Лапшин. 2018. С. 36-39.

16. Крюков О.В., Серебряков А.В. Активно-адаптивные алгоритмы управления и мониторинга автономными энергетическими комплексами // В сборнике: Пром-Инжиниринг. Труды II МНТК. ЮУрГУ. 2016. С. 286-290.
17. Vasenin A.V., Kryukov O.V., Serebryakov A.V. Adaptive control algorithms of autonomous generator complexes // В книге: Электромеханика, электротехнологии, электротехнические материалы и компоненты. Труды МКЭЭЭ-2016. 2016. С. 133-135.
18. Благодаров Д.А., Крюков О.В. Показатели энергосбережения в системах интеллектуального электропривода // В сб.: Великие реки' 2019 труды научного конгресса. 2019. С. 28-31.
19. Крюков О.В., Степанов С.Е. Новые конструкции электроприводных ГПА // В сборнике: Труды XVII Международной конференции "Электромеханика, Электротехнологии, Электротехнические Материалы и Компоненты". 2018. С. 177-180.
20. Крюков О.В., Леонов В.П., Федоров О.В. Применение микропроцессорной техники в нагружающих устройствах // Двигателестроение. 1987. № 7. С. 37-43.
21. Крюков О.В., Серебряков А.В. Экологические направления электроснабжения и задачи энергосбережения при реконструкции объектов // Электрооборудование: эксплуатация и ремонт. 2015. № 8. С. 23-33.

D.A. Blagodarov¹, Yu. M. Safonov¹, O.V. Kryukov²

NEW DESIGN AND TECHNOLOGICAL SOLUTIONS IN ENERGY-SAVING DRIVES

¹*Moscow power engineering Institute (TU), Moscow,*

²*TSN-electro LLC, Nizhny Novgorod*

An analysis of the designs of electric drives with megawatt class AC machines for integrated versions of oil and gas industry units is proposed. The main advantages of encapsulated machines in electric drives of gas pumping units are reviewed. The design features of integrated electric motors with turbochargers on the main gas pipelines of Russia are considered. The effective results of the implementation of gearless and oil-free technologies for improving the functionality, reliability, resource, energy efficiency and environmental friendliness of the most critical electromechanical systems are shown.

Keywords: electric motor, design, electric machine, rotation axis, reliability, electric drive, turbocharger.

ИНТЕЛЛЕКТУАЛИЗАЦИЯ ОБЪЕКТОВ ЭНЕРГОМАШИНОСТРОЕНИЯ

¹ ООО «ТСН-электро», г. Н. Новгород, ² НИУ «МЭИ», г. Москва

Рассмотрены современные подходы к использованию парадигмы искусственного интеллекта в вопросах управления и мониторинга различными сложными электромеханическими системами и автоматизированными электроприводами технологических установок. Предложена классификация методов и сравнительная оценка их возможностей. Приведены примеры практической реализации методологии искусственных нейронных сетей для прогнозирования технического состояния электроприводов турбокомпрессоров.

Ключевые слова: электромеханическая система, искусственные нейронные сети, электропривод, искусственный интеллект, автоматическое управление, прогнозирование.

Введение

Применение методов искусственного интеллекта (ИИ) позволяет осуществить попытку найти новые подходы к решению проблемы адекватного прогнозирования технического состояния мощных и дорогостоящих машин, на основе применения знаний и опыта экспертов, теории распознавания образов, обучения, адаптации к изменяющимся внешним условиям, принятия решения при дефиците информации, наличии шумов [1-4].

Как известно, методы ИИ разделяются на два подхода: детерминированные и эволюционные. К первым относятся системы на базе нечеткой логики [5-8]. К эволюционным методам – генетические алгоритмы и искусственные нейронные сети (ИНС) [9-12]. Сравнительные характеристики различных методов ИИ и их возможное применение при решении задач диагностирования технического состояния (ТС) электромеханических систем (ЭМС), включая синхронные двигатели (СД) [13-16] представлены в таблице 1.

*Таблица 1***Сравнительная характеристика методов искусственного интеллекта**

Метод	Особенности	Недостатки (для решаемой задачи)	Использование при оценке ТС
Нечеткая логика	Распознавание образов, принятие решений	Не могут решать задачи генеза и прогнозирования ТС	Использование совместно с другими методами ИИ на этапе решения задач классификации
Генетические алгоритмы	Универсальная оптимизация	Напрямую не применимы к задачам диагностирования	Использование совместно с другими методами ИИ на этапе настройки параметров систем
Нейронные сети	Типовые элементы, параллельная обработка, универсальная аппроксимация, классификация, распознавание образов	От выбора оптимальной структуры ИНС зависит от качества решения задачи	Подходят для решения всех типов задач диагностирования

Методы нечеткой логики могут быть эффективны на стадиях распознавания и классификации тех или иных отказов, например, в работах [17-21], но не отвечают требованиям решения задач генеза и прогнозирования состояния объекта в среднесрочной перспективе.

Перспективность моделей искусственных нейронных сетей

Искусственные нейросетевые (ИНС) модели являются универсальными вычислительными структурами, поэтому могут быть наиболее эффективны для всех необходимых задач оценки ТС ЭМС. Такие свойства ИНС как универсальная аппроксимация должны оптимальным образом решать задачи генеза. Широкие возможности по классификации и распознаванию образов позволяют выделять различные классы состояний при решении задачи оценки ТС ЭМС. Экстраполирующие свойства вычислительных нейронных структур могут обеспечить решение задач прогнозирования ТС ЭМС. Широкое применение ИНС нашли для решения задач изучения сложных процессов, где отсутствуют способы формального описания исследуемых объектов и явлений. При решении технических задач, таких, например, как:

- идентификация объектов управления,
- структурно-параметрический синтез регуляторов,
- диагностирование неисправностей, и др.

Применение НС до сих пор достаточно ограничено. Отчасти это объясняется наличием развитого математического аппарата, описывающего процессы и явления, происходящие в объектах, а также наличием большого количества методов решения указанных задач.

Несмотря на это, в последние годы наблюдается положительная динамика в плане применения ИНС для решения задач идентификации и диагностирования неисправностей технических объектов (авиационных газотурбинных двигателей, асинхронных двигателей и СД, мехатронных автономных модулей движения, различных роботов для внутритрубной диагностики магистральных нефтегазопроводов и технологических внутривысотных систем), а также проблем оперативного прогнозирования состояния оборудования с целью повышения их эффективного функционирования. Однако области применения этих исследований ограничиваются простыми объектами, не учитывают эмерджентность эксплуатационных свойств всей ЭМС. Кроме того, усложнение современных технологий требует совершенствования систем диагностирования, повышающих быстродействие при поиске отказов и оценке ТС и достоверность прогнозирования ТС, которые не могут быть обеспечены существующими методами и средствами.

Необходимо отметить, что ИНС лишены недостатков, присущих другим методам ИИ:

- им не требуется большая база знаний об объекте, как экспертным системам;
- для них не нужно описывать логику работы и функции принадлежности, как в методах нечеткой логики. ИНС имеют в себе резервы для дальнейшего развития задач оценки работоспособности. Возможна комбинация ИНС с другими методами ИИ, например, в качестве дополнительной помощи при распознавании неисправности, или на завершающей стадии в системах принятия решений (СПР), помогая человеку принять единственно верное решение.

Существует тенденция в научных исследованиях по реализации различных моделей ИИ на основе ИНС: нейронные экспертные сети, гибридные НС, лингвистические НС, нейронные семантические сети. Это подтверждает перспективность дальнейшего совершенствования теории ИНС для ЭМС.

Таким образом, на сегодняшний день ИНС являются уникальным и универсальным инструментом, который можно использовать на всех стадиях процесса диагностирования.

Аппарат ИНС позволяет реализовать программно-математическое обеспечение, позволяющее, не только выполнять однажды запрограммированную последовательность действий над строго определенными данными, но и способное обобщать вновь поступающую информацию, находить в ней закономерности, и т.д. Нейросетевое моделирование позволяет ответить на вопрос поискового прогноза ТС ЭМС "что будет, если?". Создание обучающей выборки, ее корректность, подбор обучающих параметров

одна из основных задач разработчика нейросетевой модели для оценки ТС ЭМС. Например, разработчику экспертной системы не требуется устанавливать взаимосвязи между входными данными и необходимым решением, тратя время на разнообразную статистическую обработку, подбор математического аппарата, создание и длительную проверку математических моделей. Основная задача состоит в нахождении параметров "серого ящика", оптимальных в смысле точности отображения входных обучающих векторов в выходные. Главным критерием работы нейросетевых экспертных систем является практика - многократные испытания и проверки в самых различных условиях управления.

Преимущество применения нейросетевых технологий в электромеханике и электроприводе может быть обусловлено следующими соображениями:

- возможность оценки ТС не одного элемента (двигателя, электромеханической конструкции), а функционально связанного множества элементов (элементов управления двигателей, кабеля, трансформатора и т.д.) с помощью интеллектуальных моделей. Учет прецедентности (регистрация фактов нахождения электромеханического оборудования в тех или иных ТС) в процессе эксплуатации больших ЭМС;
- быстрой параллельной обработке высокоразмерных векторов – контролируемых данных, реализующих нейросетевые модели контроля, прогнозирования ТС и поиска неисправности элементов ЭМС, с встройкой сложных программ диагностирования ТС в контур управления ЭМС;
- возможность программного обучения ИНС, при котором она приобретает "знание" об исследуемом предмете на основании сигналов, что его описывают, поэтому наученную ИНС можно применять к решению задачи без необходимости знания алгоритма, нахождения решения;
- правильно наученная ИНС имеет способность к обобщению, то есть ее можно использовать для решения задач интерполяции и экстраполяции, т.е. она способна сгенерировать правильное решение для ситуаций, которые не были предусмотрены за время обучения.

Применительно к ЭМС нейросетевые технологии способны найти применение в качестве средств распознавания (при контроле ТС ЭМС), прогнозирования и моделирования процессов развития предаварийных ситуаций и поиске неисправности. Анализ показал, что имеется необходимость построением эффективных методов комбинаторного поиска (комбинаторной оптимизации) при диагностировании состояния системы "параметры технического состояния ЭМС - внешние факторы" для класса экстремальных задач. ИНС в перечисленных системах используется в качестве самостоятельной системы представления знаний, и в практических приложениях выступает, как один из компонентов системы управления либо модуля принятия решений и может быть принята для оценки ТС ЭМС.

Техническая реализация систем искусственного интеллекта на базе ИНС

В настоящее время, наиболее массовым и доступным направлением использования ИНС является моделирование нейронных сетей на обычных компьютерах, прежде всего персональных, которые можно разместить в операторских кабинах. Априорная информация об алгоритмических возможностях нейросетевых архитектур, таких как, отсутствие вычислительных сложностей при росте размерности задачи распознавания, способность разделять в пространстве признаков выпуклые и несвязанные области, а также высокие аппроксимационные возможности нейросетевых моделей позволяет сделать заключение о перспективности и доступности нейросетевых технологий в диагностировании ТС ЭМС. В современной литературе по оценке ТС ЭМС практически отсутствуют основы теории применения ИНС для целей оценки ТС, прогнозирования ТС

и поиска неисправности. Таким образом, высокая размерность пространства состояний СД в комплексе с высокими требованиями к оперативности диагностических решений в контуре управления режимами ЭМС, требуют применения нейросетевых технологий.

В настоящее время исследованиями в этой области занимается свыше 300 зарубежных фирм, причем число их постоянно увеличивается. Наблюдается тенденция перехода от программной к программно-аппаратной реализации нейросетевых алгоритмов со значительным увеличением производства нейрочипов и нейроконтроллеров. Существенно увеличилось число военных разработок, ориентированных на создание сверхбыстрых, «умных» суперЭВМ. Нейросетевые алгоритмы являются попыткой моделирования способности человеческого мышления, в частности способности, обучаться и решать задачи распознавания по прецедентам. Доступность и возросшие вычислительные возможности современных компьютеров привели к широкому распространению программ, использующих принципы нейросетевой обработки данных на последовательных ЭВМ. Этот подход не использует преимуществ параллелизма, присущих нейро-вычислениям и ориентируется исключительно на возможность с помощью моделей ИНС решать слабоформализуемые задачи, так или иначе связанных с обработкой образов.

Методология разработки ИНС ТС СД

Решение проблемы разработки методологических основ построения методов диагностирования сложных ЭМС на основе нейросетевых технологий связано с обоснованием и экспериментальной проверкой в условиях действующего производства их информационного, математического и программного обеспечения. При этом необходимо обеспечить повышение эффективности процессов диагностирования состояний сложных ЭМС путем научного обоснования и практического подтверждения методологии применения новых методов, моделей, алгоритмов и программного обеспечения на базе распространенных архитектур ИНС. В соответствии с этим последовательный алгоритм решения следующий:

1. Анализ существующих методов контроля и прогнозирования технических состояний электромеханического оборудования и обоснования целесообразности создания интеллектуальной системы диагностирования сложных ЭМС на базе распространенных архитектур нейронных сетей.
2. Разработка концепции совершенствования систем диагностирования сложных электромеханических систем с учетом применения ИНС в условиях нестационарности и многофакторности воздействий внешней среды.
3. Создание методов и моделей контроля технических состояний сложных ЭМС в нейросетевом базисе при иерархической системе распознавания.
4. Разработка методов, моделей и алгоритмов для решения задачи прогнозирования технических состояния электромеханических систем на основе экстраполяции в пространстве возможных прецедентов их эксплуатации.
5. Синтез алгоритмического и программного обеспечения и создание аппаратного комплекса технического диагностирования электромеханических систем в режиме эксплуатации.
6. Экспериментальная проверка, разработанных методов, моделей и алгоритмов диагностирования состояний электромеханических систем.

Перспективные направления развития технологий ИНС для электроприводов

Анализ состояния вопроса в области теории и практики диагностирования электромеханических систем показал следующее:

1. Новые требования оценки эффективности функционирования состояний ЭМС связаны главным образом с переходом к получению и использованию результатов не отдельных измерений, а потоков информации о десятках однородных или разнородных

величин, большая часть из которых может быть недоступной для прямых измерений. Использование ЭМС в опасных условиях функционирования делают участие человека-оператора вообще недопустимо. Таким образом, при диагностировании ЭМС необходимо создание новых средств, позволяющих исключить человека из процесса сбора и обработки интенсивных потоков измерительной информации и принятия решений о состоянии электромеханической системы.

2. В результате анализа состояния исследований в областях оценки ТС сложных объектов установлено, что наиболее перспективным является направление с использованием ИНС, которые позволяют в едином методологическом базисе формулировать задачи оптимального управления, и в реальном масштабе времени решать задачи контроля ТС, поискового прогнозирования ТС и поиска неисправности ЭМС, а современная компьютерная элементная база, позволяет реализовать нейросетевой вычислительный базис диагностических моделей ЭМС, с учетом жестких ограничений на массовые, габаритные и энергетические характеристики средств обработки, в неблагоприятных условиях применения ЭМС.

3. Изучение состояния исследований в области создания систем диагностирования ТС ЭМС выявило состав подсистем, при которых эта оценка максимальна по достоверности и полноте: контроля готовности технологического оборудования; оперативного контроля и прогнозирования ЭМС; контроля функционирования ЭМС и поиска неисправностей ЭМС.

4. На основе анализа состояния вопроса представлено описание решения второй задачи диагностирования ЭМС - прогнозирования ее технического состояния в нейросетевом алгоритмическом базисе.

5. Классификация методов поиска отказов ЭМС выявила три основных направления:

- проверки элементов ЭМС последовательно с анализом ТС после каждой проверки;
- проверки групп элементов ЭМС с анализом ТС с гибкой программой поиска;
- проверки групп элементов ЭМС с анализом ТС после проведения полной группы проверок, обеспечивающей однозначное определение неисправного элемента.

Таким образом, проведенная классификация позволяет произвести количественную оценку методов, а, следовательно, их сравнение и выработку действенных практических рекомендаций. Синтез оптимальных поисковых процедур позволяет совершенствовать систему поиска и определения причин неисправностей, по двум направлениям разработок: выбор оптимального объема контрольных проверок элементов ЭМС и разработка алгоритма выбора оптимальной последовательности проверок. Частично решить проблемы, связанные с экспоненциальным ростом числа возможных комбинаций в последовательности проверок, позволяют параллельные реализации методов поиска мест, и причин неисправностей ЭМС.

6. Предлагается для диагностирования состояний ЭМС использовать распространенные архитектуры НС, которые обладают рядом важных свойств:

- решать плохо формализуемые традиционными математическими методами задачи, к числу которых относится идентификация ТС ЭМС, причем решать их в режиме реального времени за счет высокой скорости работы ИНС, при дефиците измерительной информации и наличии шумов в измерениях, большом количестве измеряемых параметров;
- в процессе своей работы ИНС способна самообучаться в соответствии с изменяющимися в процессе эксплуатации ЭМС параметрами;
- позволяют учесть не только изменения технологических режимов работы, но и изменения состояния ЭМС;
- обладают большим быстродействием по сравнению с моделями, построенными на основе дифференциальных или алгебраических уравнений.

Выводы

Анализируя перспективы развития и использования современных систем автоматизации для задач машиностроения в России, можно прогнозировать рост потребности в них в связи с модернизацией и созданием нового технологического оборудования. Применение интеллектуальных модулей и микропроцессорных средств с ИТ- алгоритмами позволяет:

- идентифицировать параметры объектов, осуществлять самонастройку регуляторов, адаптацию к изменению параметров, мониторинг, что делает систему *контролируемой*;
- реализовать новые способы управления и прогнозирования на базе нейронных сетей и генетических алгоритмов, что делает систему *обучаемой*;
- использовать информационный обмен и сетевые технологии, обеспечивая связь между модулями и с системами управления высокого уровня, что делает ее *интеллектуальной*.

Библиографический список

1. Васенин А.Б., Степанов С.Е., Крюков О.В. Методология и средства оперативного мониторинга электродвигателей на КС // Контроль. Диагностика. 2019. № 11. С. 52-58.
2. Kryukov O.V., Blagodarov D.A., Dulnev N.N., Safonov Y.M., Fedortsov N.N., Kostin A.A. Intelligent control of electric machine drive systems // В сборнике: 2018 10th International Conference on Electrical Power Drive Systems, ICEPDS 2018 - Conference Proceedings 10. 2018.
3. Туганов Р.Б., Благодаров Д.А., Крюков О.В. Перспективы автоматизации объектов машиностроения // В сборнике: Информационные системы и технологии ИСТ-2020. Сборник материалов XXVI МНТК. НГТУ им. Р.Е. Алексеева. 2020. С. 323-330.
4. Крюков О.В., Серебряков А.В. Активно-адаптивные алгоритмы управления и мониторинга автономными энергетическими комплексами // В сборнике: Пром-Инжиниринг. Труды II МНТК. ЮУрГУ. 2016. С. 286-290.
5. Мещеряков В.Н., Ласточкин Д.В., Крюков О.В. Приложения теории нечетких множеств для обработки данных и задач прогнозирования в системах АЭП // В сборнике: Современные сложные системы управления. Материалы XII МНПК. 2017. С. 153-158.
6. Крюков О.В. Моделирование и микропроцессорная реализация электромеханических систем // Электротехника: сетевой электронный научный журнал. 2015. №3. С. 55-61.
7. Крюков О.В., Леонов В.П., Федоров О.В. Применение микропроцессорной техники в нагружающих устройствах // Двигателестроение. 1987. № 7. С. 37-43.
8. Vasenin A.B., Kryukov O.V., Serebryakov A.V. Adaptive control algorithm so fautomous generator complexes // В книге: Электромеханика, электротехнологии, электротехнические материалы и компоненты. Труды МКЭЭЭ-2016. 2016. С. 133-135.
9. Мещеряков В.Н., Крюков О.В., Туганов Р.Б. Применение методов искусственного интеллекта для управления и мониторинга электромеханических систем и электроприводов // Автоматизация и ИТ в энергетике. 2020. № 3 (128). С. 10-16.
10. Крюков О.В. Алгоритмы быстрого преобразования Уолша в микропроцессорных системах управления электроприводом // Известия ВУЗов. Электромеханика. 2005. № 4. С. 39-44.
11. Васенин А.Б., Крюков О.В. Проектирование электромеханической части и систем управления энергетических установок газотранспортных потребителей // Известия Тульского гос. ун-та. Технические науки. 2011. № 5-1. С. 47-51.

12. Степанов С.Е., Васенин А.Б., Крюков О.В. Вейвлет-преобразование в задачах анализа и синтеза АЭП // В сборнике: Информационные системы и технологии ИСТ-2020. XXVI МНТК. НГТУ им. Р.Е. Алексеева. 2020. С. 307-314.
13. Крюков О.В. Мониторинг условий эксплуатации электродвигателей газоперекачивающих агрегатов // Контроль. Диагностика. 2016. № 12. С. 50-58.
14. Крюков О.В., Степанов С.Е. Модернизация систем управления ЭГПА в условиях действующих компрессорных станций // В сборнике: Проблемы автоматизации и управления в технических системах. МНТК под редакцией М.А. Щербакова. 2013. С. 29-32.
15. Крюков О.В., Мещеряков В.Н., Гуляев И.В. Электроприводы на основе машины двойного питания и АВК с преобразователями в цепях статора и ротора // Саранск, 2020.
16. Серебряков А.В., Крюков О.В. Новые возможности гибридных энергетических установок на платформе SMARTGRID // В сборнике: Актуальные проблемы электроэнергетики. Материалы ВНТК. Нижний Новгород, 2012. С. 135-137.
17. Крюков О.В., Серебряков А.В. Экологические направления электроснабжения и задачи энергосбережения при реконструкции объектов // Электрооборудование: эксплуатация и ремонт. 2015. № 8. С. 23-33.
18. Васенин А.Б., Крюков О.В. Энергоэффективные и экологичные установки воздушного охлаждения // В сборнике: ВЕЛИКИЕ РЕКИ' 2017 труды научного конгресса 19-го МНПФ. НГАСУ. 2017. С. 93-96.
19. Благодаров Д.А., Крюков О.В. Показатели энергосбережения в системах интеллектуального электропривода // В сб.: Великие реки' 2019 труды научного конгресса. 2019. С. 28-31.
20. Kryukov O.V., Gulyaev I.V., Teplukhov D.Y. Method for stabilizing the operation of synchronous machines using a virtual load sensor // Russian Electrical Engineering. 2019. № 7. С. 473-478.
21. Благодаров Д.А., Сафонов Ю.М., Крюков О.В. Энергосберегающие системы автоматизированного электропривода турбокомпрессоров // В сборнике: Информационные системы и технологии ИСТ-2020. XXVI МНТК. НГТУ им. Р.Е. Алексеева. 2020. С. 269-276.

S. V. Vorobyov¹, O. V. Kryukov¹, D. A. Blagodarov²

INTELLECTUALIZATION OF POWER ENGINEERING FACILITIES

¹ООО "TSN-Elektro", N. Novgorod, ²National research UNIVERSITY "MPEI", Moscow

Modern approaches to the use of the artificial intelligence paradigm in the management and monitoring of various complex electromechanical systems and automated electric drives of technological installations are considered. A classification of methods and a comparative assessment of their capabilities are proposed. Examples of practical implementation of the methodology of artificial neural networks for predicting the technical condition of electric turbocharger drives are given.

Keywords: electromechanical system, artificial neural networks, electric drive, artificial intelligence, automatic control, forecasting.

СИСТЕМА АВТОМАТИЗАЦИИ ОБЪЕКТОВ ТЭК И МАГИСТРАЛЬНЫХ НЕФТЕПРОВОДОВ

¹ Нижегородский государственный технический университет им. Р.Е. Алексеева,

² ООО «ТСН-электро», г. Нижний Новгород

Представлены результаты создания системы автоматизированного диспетчерского контроля и управления технологическими объектами нефтепроводов в рамках комплекса работ по проектированию системы магистральных нефтяных трубопроводов Республики Казахстан. Рассмотрены инновационные решения, обеспечивающие автоматическое управление всеми объектами нефтеперекачивающих станций, повышение уровня безопасности и энергоэффективности эксплуатации нефтепроводов и обеспечение качества транспортируемых нефтепродуктов. Предложено описание особенностей функционирования структурной схемы автоматизированной системы управления технологическим процессом транспорта нефти на основе концепции открытых международных стандартов всех уровней IT-взаимодействий компонентов.

Ключевые слова: автоматизированная система управления, мониторинг, контроль, магистральный трубопровод, нефтеперекачивающая станция.

Введение

Стабильное развитие топливно-энергетического комплекса России является одним из стратегических направлений в экономике России и Казахстана [1-4]. В этой связи особую актуальность приобретает дальнейшее развитие нефтетранспортной инфраструктуры для обеспечения роста экспорта нефти, а также ее поставок на внутренние рынки. Для этого, например, компания ЗАО «КазТрансОйл» как основной оператор поставок казахстанской нефти на экспорт выполняет целый комплекс работ:

- оптимизацию поставок нефти,
- объединение системы нефтяных трубопроводов в единую технологическую систему,
- обеспечение взаимодействия с операторами, потребителями и поставщиками нефти,
- организацию менеджмента качества нефти и воздействия на окружающую среду в соответствии с требованием международных стандартов ISO.

Поставленные задачи решаются на основе современных информационных технологий [5-9], которые уже успешно апробированы в смежных отраслях промышленности России [10-14]. В частности, по заказу компании АВВ (Германия) выполнен комплекс проектных работ по созданию АСУ ТП нефтеперекачивающих станций (НПС) магистральных нефтепроводов (МН), проходящих по территории Республики Казахстан. Состав работ определялся в соответствии с ГОСТ 34.601-90 "АСУ. Стадии создания". Необходимость в комплексном обследовании объектов с выработкой предложений по подготовке технологических процессов на объектах к автоматизации обусловило разработку проекта и рабочей документации, включая информационно-алгоритмическое обеспечение, проведение авторского надзора и подготовку документации "Как построено" [15-18].

В данном проекте задача создания системы автоматизированного диспетчерского контроля и управления нефтепроводами и технологическими объектами ЗАО «КазТрансОйл» должна обеспечить автоматизированное управление всеми технологическими объектами, повышение уровня безопасности эксплуатации нефтепроводов, обеспечение качества нефти в соответствии с требованиями стандарта ISO, мониторинг технологического оборудования, обнаружение и локализацию утечек [19-23]. Объекты, для которых выполнялось проектирование АСУ ТП, выделены на рис. 1.

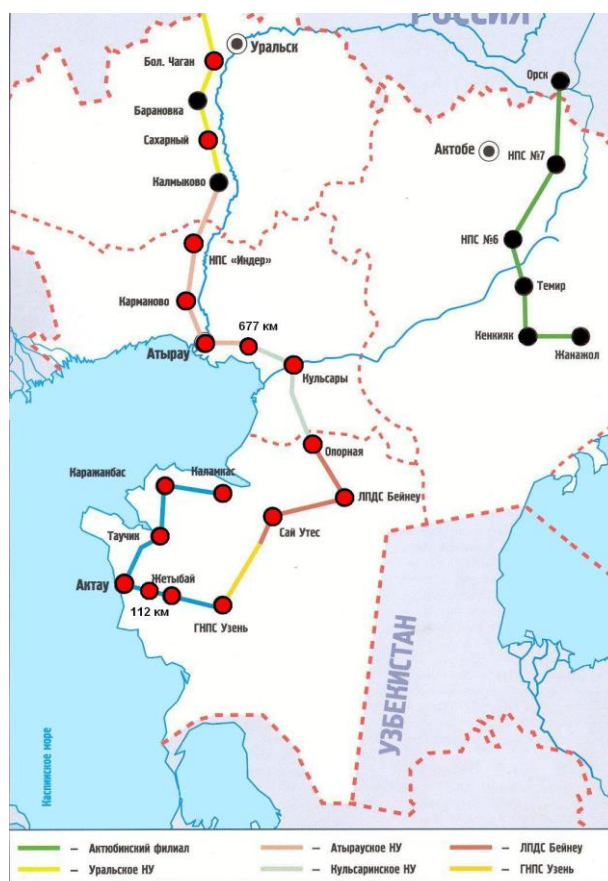


Рис. 1. Схема трубопроводов Западного и Актыбинского филиалов «КазТрансОйл»

Основными объектами автоматизации явились головные нефтеперекачивающие станции (ГНПС), линейные НПС и станции подогрева нефти (СПН).

Особенности проекта автоматизации

В составе вышеназванных объектов вошли магистральные насосные, подпорные насосные, резервуарные парки, наливные/сливные железнодорожные эстакады и танкерные причалы, узлы пуска и приема устройств очистки и диагностики трубопроводов, пункты подогрева нефти, узлы регулирования давления нефти с запорной арматурой, узлы учета нефти, вспомогательные системы, обеспечивающие функционирование НПС.

При разработке проекта была решена задача использования однородных ПТС на каждом из уровней управления. Для автоматизации НПС, нефтехранилищ и СПН были использованы программно-технические средства Freelance 2000 (поставки АВВ AutomationSystemsGmbH), которые обеспечивают высокий уровень масштабирования, резервирование, диагностирование до уровня канала, простоту обслуживания, функции мониторинга и контроля.

Функциональность АСУ ТП обеспечила создание автоматизированной многоуровневой системы на базе современных ПТС, реализацию принципа управления «сверху» путем приема в автоматическом режиме плановых заданий и установок работы НПС с вышестоящего уровня, обеспечение непрерывного контроля работы основного технологического оборудования и систем жизнеобеспечения станции, оптимизацию режимов работы станций, повышение технико-экономических показателей за счет своевременного определения, локализации и устранения утечек, а также сокращения простоев за счет определения оптимального режима перекачки.

Особенностью проектируемой СА являлось внедрение ее на существующих, действующих объектах (НПС, СПН), что предъявляло соответствующие требования к основной технологии и системам жизнеобеспечения. При этом заказчиком были поставлены задачи по замене устаревшего и изношенного оборудования КИП,

исполнительных устройств и механизмов, систем локальной автоматики на современные микропроцессорные системы с интеграцией в создаваемую АСУ ТП НПС МН (рис. 2). В составе работ был проведен анализ наличия введенных в действие микропроцессорных систем и современной датчиковой аппаратуры, по результатам которого выработаны решения по их интеграции в АСУ ТП.

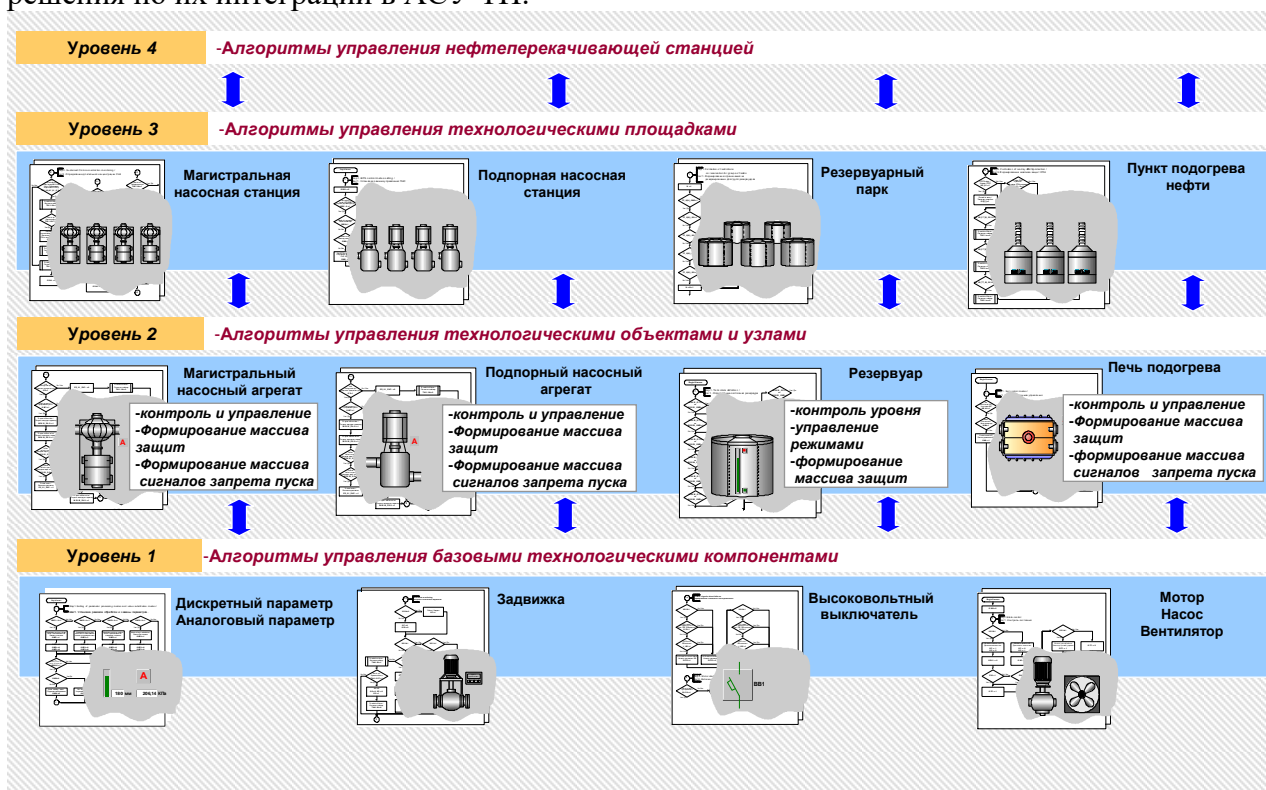


Рис. 2. Функциональная структура алгоритмического обеспечения АСУ ТП

В ходе детального технического обследования определен объем реконструкции морально и физически устаревшего технологического и электротехнического оборудования, а также части трубопроводов. Разработаны предложения по подготовке объектов к внедрению новой современной системы АСУ ТП.

Учитывая динамику изменений технологических задач и развития инфраструктуры объектов, АСУ ТП создавалась изначально с открытой архитектурой, позволяющей дополнять систему при вводе в строй новых технологических объектов или участков магистральных нефтепроводов, либо заменять морально устаревшие компоненты системы при минимальных затратах на стыковку с существующим комплексом ПТС.

Система диспетчерского контроля и управления нефтепроводами и технологическими объектами ЗАО «КазТрансОйл» является многоуровневой. Благодаря тому, что технологическое оборудование во многом подобно, были созданы стандартные алгоритмические модули, которые адаптировались для конкретных объектов автоматизации. Таким образом, однородность создаваемой системы была обеспечена не только на аппаратном уровне технических средств, но и на алгоритмическом и программном уровне.

Проиллюстрируем проектные решения на примере АСУ ТП головной НПС (4-й уровень диспетчерского управления), где представлены практически все виды ПТС, примененные в проекте. Структурная схема АСУ ТП ГНПС, представлена на рис. 3.

Структура комплекса программно-технических средств АСУ ТП ГНПС (как и всей АСУ ТП НПС МН ЗАО «КазТрансОйл») базируется на следующих основных принципах построения АСУ ТП:

- централизованный, иерархический контроль и управление технологическими объектами и магистральными нефтепроводами;
- открытая архитектура IT-взаимодействия компонентов АСУ;

- распределенная структура подсистемы сбора и обработки;
- оптимизация распределения функций сбора информации, контроля, управления на базе объектно-ориентированного подхода;
- простота ТО и высокая степень готовности ПТС;
- короткое время восстановления системы;
- самодиагностика и выборочное дублирование или резервирование компонентов комплекса ПТС.

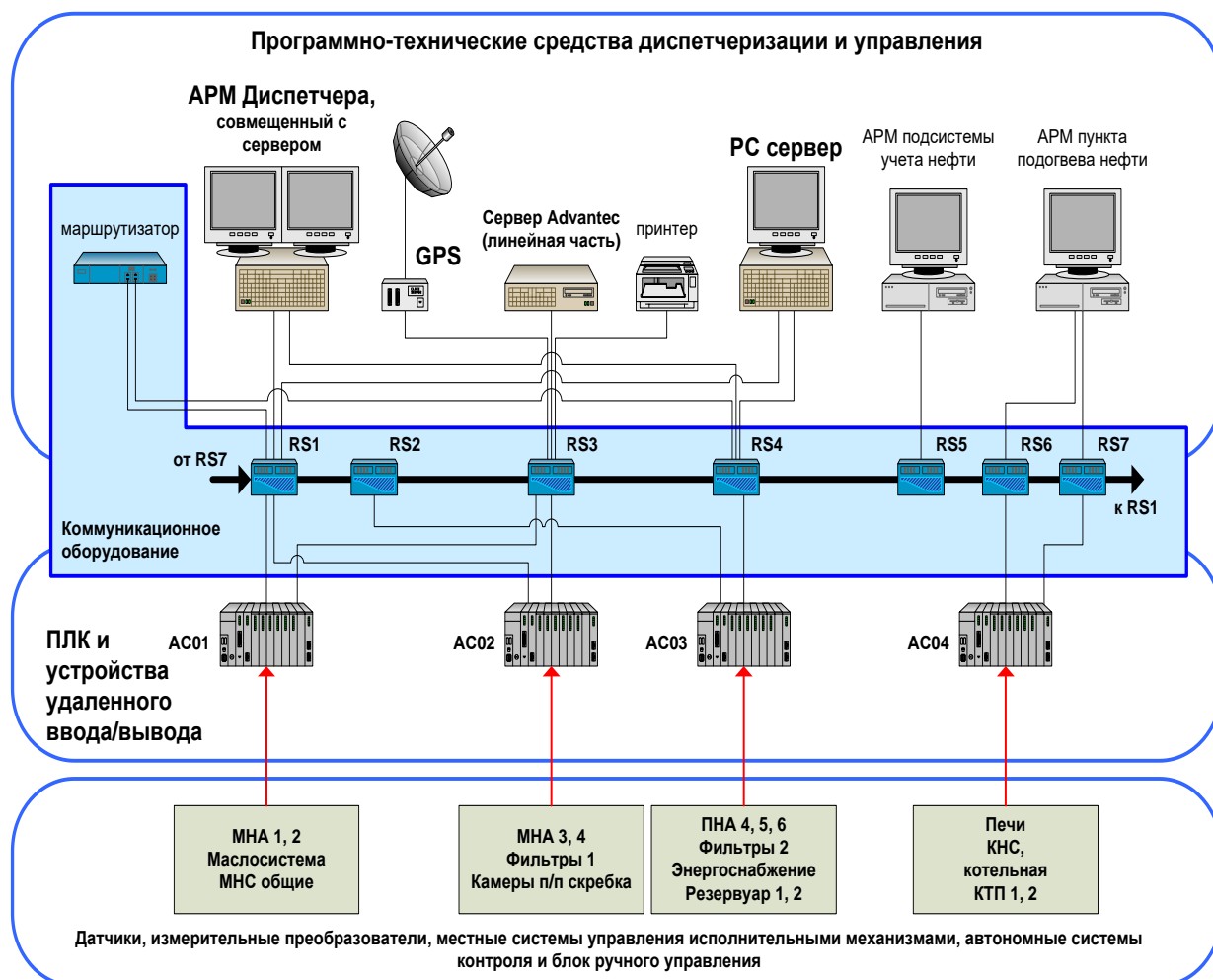


Рис. 3. Структурная схема АСУ ТП ГНПС

Принятая концепция построения структуры комплекса ПТС предусматривает применение открытых международных стандартов для всех уровней IT-взаимодействия компонентов СА. Таким образом, обеспечивается будущее расширение и модернизация АСУ ТП.

Принятая концепция построения структуры комплекса ПТС определяет взаимосвязь между устройствами и уровнями ПТК с максимально возможным приближением к функционально-групповому принципу построения технологического объекта управления (ТОУ).

Структурная схема технических средств АСУ ТП ГНПС разработана по принципам многоуровневой, иерархической информационно-управляющей системы и имеет 3 уровневую иерархию: ПТС диспетчеризации и управления; ПЛК AC800F и устройства удаленного ввода/вывода S800 и Датчики, измерительные преобразователи, местные САУ исполнительными механизмами, автономные системы контроля и блок ручного управления.

Структура реализована в виде распределенной по технологическим объектам системы, с целью образования локальных децентрализованных структур.

ПТС на уровне ДП ГНПС могут функционировать в автоматизированных режимах: в полном составе; с включением резервных элементов системы и не в полном составе.

Первый режим является основным, два последних – резервными для обеспечения живучести системы при возникновении внештатных или аварийных ситуаций. Для повышения надежности задачи диагностики комплекса ПТС выполняются на всех уровнях с использованием функций самодиагностики. На уровне контроллеров формируются диагностические признаки состояния, измерительных каналов, модулей S800, источников питания. На уровне сервера SCADA диагностируются состояния контроллера и сети Ethernet.

Аппаратура системы диагностируется автоматически в процессе работы. При необходимости может быть организован режим контроля системы для более глубокой диагностики, в этом случае контроль должен запускаться оператором с рабочих мест системы. Данные о неисправности аппаратуры вводятся в базу данных, и отображаются на экране с указанием отказавшего узла системы с точностью до блока и регистрируются в протоколе событий.

В АСУ ТП ГНПС реализованы функции: коммуникационные, информационные, дистанционного управления, формирования отчетов, настройки компонент АСУ ТП.

В число автоматизируемых ТОУ входят СПН, предназначенные для подогрева транспортируемой высоковязкой нефти. При перекачке с подогревом СА должна обеспечивать выполнение соответствующих расчетов и определение условий работы нефтепровода с наименьшим энергопотреблением и поддержание оптимального режима эксплуатации. При этом осуществляется контроль за состоянием с дистанционным управлением печами подогрева, задвижками технологических нефтепроводов, системой внутренней циркуляцией нефти, системой топлив обеспечения горелок печей, системой стационарного пожаротушения.

САУ обеспечивают автоматические защиты от недопустимых режимов работы ТП, в т.ч. защиту от перелива нефти в резервуары при достижении в них максимального (аварийного) уровня нефти и переключение потока нефти в специально выделенные емкости.

Одной из важной составляющей внедряемой SCADA является возможность обнаружения утечек (COY) для определения дефектов, возникших под влиянием природных факторов или хищений из нефтепровода и позволяющая исключить или значительно сократить катастрофические экологические последствия.

Выводы

Таким образом, комплекс проектных работ по созданию АСУ ТП нефтеперекачивающих станций магистральных нефтепроводов, проходящих по территории Республики Казахстан, позволил создать интегрированную систему диспетчерского контроля и управления трубопроводами и технологическими объектами. Проектные решения позволили обеспечить выполнение современных требований к созданию подобных больших автоматизированных систем: функциональность, надежность, устойчивость, однородность, совместимость.

Библиографический список

1. Крюков О.В., Мещеряков В.Н. Электрооборудование систем электроснабжения автономных объектов нефтегазопроводов // Электрооборудование: эксплуатация и ремонт. 2020. №7. С. 28-36.
2. Васенин А.Б., Крюков О.В. Средства оптимизации управления электроприводов магистральных газопроводов // В сборнике: Труды XVII Международной конференции "Электромеханика, Электротехнологии, Электротехнические Компоненты". 2018. С. 180-183.

3. Крюков О.В., Серебряков А.В. Экологические направления электроснабжения и задачи энергосбережения при реконструкции объектов // Электрооборудование: эксплуатация и ремонт. 2015. № 8. С. 23-33.
4. Крюков О.В. Комплексный анализ условий эксплуатации электродвигателей ГПА // Компрессорная техника и пневматика. 2013. № 4. С. 14-19.
5. Khlynin A.S., Kryukov O.V., Stepanov S.E. Energy optimization of gas compressor units electric drives // В книге: Electromechanics, Electrotechnology, Electromaterials and Components. Труды МКЭЭЭ-2016. 2016. С. 137-139.
6. Васенин А.Б., Крюков О.В. Проектирование электромеханической части и систем управления энергетических установок газотранспортных потребителей // Известия Тульского гос. ун-та. Технические науки. 2011. № 5-1. С. 47-51.
7. Крюков О.В. Современные системы электроснабжения с энергетическими установками вдоль трассовых потребителей газопроводов // В сборнике: Актуальные проблемы электроэнергетики. Сборник научно-технических статей XXXII региональной научно-технической конференции, посвященной 75-летию проф. С.В. Хватова. Нижний Новгород, 2013. С. 61-69.
8. Мещеряков В.Н., Ласточкин Д.В., Крюков О.В. Приложения теории нечетких множеств для обработки данных и задач прогнозирования в системах АЭП // В сборнике: Современные сложные системы управления. Материалы XII МНПК. 2017. С. 153-158.
9. Крюков О.В., Серебряков А.В. Активно-адаптивные алгоритмы управления и мониторинга автономными энергетическими комплексами // В сборнике: Пром-Инжиниринг. Труды II МНТК. ЮУрГУ. 2016. С. 286-290.
10. Vasenin A. V., Kryukov O. V., Serebryakov A. V. Adaptive control algorithms of autonomous generator complexes // В книге: Электромеханика, электротехнологии, электротехнические материалы и компоненты. Труды МКЭЭЭ-2016. 2016. С. 133-135.
11. Васенин А.Б., Крюков О.В. Энергоэффективные и экологичные установки воздушного охлаждения // В сборнике: ВЕЛИКИЕ РЕКИ' 2017 труды научного конгресса 19-го МНПФ. НГАСУ. 2017. С. 93-96.
12. Крюков О.В., Степанов С.Е. Модернизация систем управления ЭГПА в условиях действующих компрессорных станций // В сборнике: Проблемы автоматизации и управления в технических системах. МНТК под редакцией М. А. Щербакова. 2013. С. 29-32.
13. Kryukov O.V., Gulyaev I.V., Teplukhov D.Y. Method for stabilizing the operation of synchronous machines using a virtual load sensor // Russian Electrical Engineering. 2019. Т. 90. № 7. С. 473-478.
14. Крюков О.В., Мещеряков В.Н., Гуляев И.В. Электроприводы на основе машины двойного питания и асинхронного вентильного каскада с преобразователями в цепях статора и ротора // Саранск, 2020.
15. Крюков О.В. Оптимальное управление технологическим процессом магистрального транспорта газа // В сборнике: XII Всероссийское совещание по проблемам управления ВСПУ-2014. Институт проблем управления им. В. А. Трапезникова РАН. 2014. С.4602-4613.
16. Kryukov O.V., Blagodarov D.A., Dulnev N.N., Safonov Y.M., Fedortsov N.N. Intelligent control of electric machine drive systems // В сборнике: 2018 10th International Conference on Electrical Power Drive Systems, ICEPDS 2018 - Conference Proceedings 10. 2018. С. 8571670.
17. Крюков О.В. Алгоритмы быстрого преобразования Уолша в микропроцессорных системах управления электроприводом // Известия ВУЗов. Электромеханика. 2005. № 4. С. 39-44.

18. Степанов С.Е., Крюков О.В. Опыт комплексной автоматизации объектов магистрального газопровода "ЯМАЛ-ЕВРОПА" // В сборнике: Информационные системы и технологии ИСТ-2018 XXIV МНТК к 100-летию Нижегородской радиолaborатории. 2018. С. 732-738.
19. Крюков О.В. Мониторинг условий эксплуатации электродвигателей газоперекачивающих агрегатов // Контроль. Диагностика. 2016. № 12. С. 50-58.
20. Воронков В.И., Рубцова И.Е., Крюков О.В. Электроснабжение и электрооборудование линейных потребителей МГ // Газовая промышленность. 2010. № 3. С. 32-37.
21. Крюков О.В. Моделирование и микропроцессорная реализация электромеханических систем // Электротехника: сетевой электронный научный журнал. 2015. №3. С. 55-61.
22. Васенин А.Б., Степанов С.Е., Крюков О.В. Методология и средства оперативного мониторинга электродвигателей на КС // Контроль. Диагностика. 2019. № 11. С. 52-58.
23. Хлынин А.С., Степанов С.Е., Крюков О.В. Нейронные сети мониторинга и прогнозирования технического состояния электропривода ГПА // В сборнике: Информационные системы и технологии ИСТ-2020 XXVI МНТК. НГТУ им. Р.Е. Алексеева. 2020. С. 331-338.

E.V. Bychkov¹, V.G. Titov¹, O.V. Kryukov²

AUTOMATION SYSTEM FOR FUEL AND ENERGY COMPLEX FACILITIES AND OIL TRUNK PIPELINES

¹ *Nizhny Novgorod State Technical University n. a. R.E. Alekseev,*

² *LLC "TSN-electro", Nizhny Novgorod*

The paper presents the results of creating computer-aided dispatch and control of technological objects pipelines as part of a complex of works on the design of the system of main oil pipelines of the Republic of Kazakhstan. Innovative solutions that provide automatic control of all objects of oil pumping stations, increase the level of safety and energy efficiency of the operation of oil pipelines and ensure the quality of transported petroleum products are considered. A description of the features of the functioning of the block diagram of the automated control system for the technological process of oil transportation based on the concept of open international standards at all levels of IT interaction of components is proposed.

Keywords: automated control system, monitoring, control, main pipeline, oil pumping station.

МЕТОДЫ “ON-LINE” ПРОГНОЗА СОСТОЯНИЯ ТРУБОПРОВОДОВ*ООО «Газпром проектирование», г. Нижний Новгород*

Представлены результаты исследовательских испытаний приборов неразрушающего контроля применительно к задаче оценки размеров колоний стресс-коррозионных трещин на поверхности труб, бывших в эксплуатации в составе магистральных газопроводов, отличных по типоразмеру, свойствам и химическому составу. Установлены закономерности между показаниями приборов, реализующих ультразвуковые, вихретоковые, электропотенциальные методы контроля и фактическими размерами трещин. Достоверность результатов исследований и выводов подтверждается применением стандартизированных методов оптической электронной микроскопии и рентгеновской томографии при верификации размеров трещин в местах максимальных показаний приборов.

Ключевые слова: магистральный газопровод, поверхность труб, методы и приборы контроля трещин, вихретоковый и электропотенциальный методы мониторинга.

Введение

Для большинства газопроводов, построенных во второй половине прошлого века, характерно образование и постепенное развитие многочисленных эксплуатационных повреждений поверхности труб в местах доступа коррозионной среды [1-4]. Общее число таких повреждений труб может достигать нескольких тысяч на один километр магистрального газопровода (МГ), при этом большинство из них имеют незначительные размеры и не оказывают существенного влияния на работоспособность газопроводов.

Общий износ МГ и экономические вызовы выдвигают на первый план помимо задач модернизации основных технологических агрегатов [5-9] планирование мероприятий по диагностированию труб и их ремонту [10-12]. Поэтому проводятся системные исследования технологии ремонта в трассовых условиях протяженных участков МГ со стресс-коррозионными повреждениями. Суть технологии заключается в возможности продолжения эксплуатации труб с дефектами коррозионного растрескивания под напряжением (КРН), глубина которых не превышает 10% от толщины стенки, а также профилактики их дальнейшего роста за счет покрытий, содержащих ингибирующие композиции [13-15].

Важнейшим аспектом технологии ремонта является классификация дефектов по степени опасности на основании результатов приборного контроля труб. На сегодняшний день, процесс определения глубины дефектов КРН представляет собой комплекс трудоемких операций, заключающихся в локализации дефектной зоны вихретоковым методом контроля, проявлении дефектов средствами магнитопорошковой дефектоскопии, контролируемой шлифовки металла в зонах с максимальными показаниями вихретоковых дефектоскопов, последующим контролем наличия остаточных трещин на вышлифованной поверхности и толщинометрией с применением контактных ультразвуковых приборов [16-19].

В этой связи актуальна разработка методов и алгоритмов обработки и оценки данных приборов НК, получаемых при контроле труб в трассовых условиях с дефектами КРН, развитие технологии НК поверхности труб для определения глубины дефектов КРН, а также разработку методической основы для обеспечения требуемой точности, производительности и экономической эффективности технического диагностирования при ремонте МГ.

Методика эксперимента

Установление возможностей применения средств НК для оценки глубины дефектов КРН было выполнено с применением 12 портативных приборов общего назначения, реализующих вихретоковый, ультразвуковой и электропотенциальный методы контроля.

При вихретоковом контроле дефектов КРН использовались дефектоскопы с возможностью работы в частотном диапазоне от 10 кГц до 1,5 МГц, с дифференциальными накладными вихретоковыми преобразователями (ВТП) трансформаторного типа. В зависимости от зоны эффективного контроля ВТП условно подразделялись на группы с диаметрами обмотки $D \leq 3$ мм и ВТП с $D > 4$ мм.

Технические средства, реализующие электропотенциальные методы НК, были представлены портативными трещиномерами с рабочей частотой тока порядка 1 кГц и четырехэлектродными (два токовых и два потенциальных электрода) датчиками линейного «1x4» и квадратного «2x2» исполнения.

При ультразвуковом контроле параметров дефектов КРН применяли дефектоскопы в ручном режиме контроля, имеющие возможность оценки размеров дефекта, запоминания и документирования параметров настроек и результатов контроля. Ультразвуковые дефектоскопы с отдельно совмещенным пьезоэлектрическим преобразователем, с частотой 5 МГц и регулировкой усиления 85 дБ с дискретностью 1 дБ. Ультразвуковой дефектоскоп-томограф с цифровой (вычислительной) фокусировкой антенной решетки и томографической обработкой данных, позволяющих получать в том числе отображение образов вертикально-ориентированных дефектов. Для работы с дефектоскопом-томографом применялась фазированная антенная решетка поперечных волн с центральной рабочей частотой 7,5 МГц и сектором обзора от 35° до 80°.

Для проведения исследований приборов НК было подготовлено 60 образцов из труб бывших в эксплуатации в составе МГ, содержащих дефекты КРН различной глубины и плотности трещин в колонии. Толщины стенок образцов варьировались от 9 до 18,7 мм, что обеспечило учет основного сортамента труб входящих в состав МГ. Образцы были изготовлены из сталей классов прочности Х60-Х70, наиболее подверженных КРН, широкой номенклатуры марок сталей (09Г2, 09Г2С, 10Г2, 17ГС, 17Г2С, 20Г2), что позволило учесть в исследованиях различные вариации электрофизических свойств и структурных неоднородностей металла, обусловленных металлургическим качеством и технологией изготовления труб [20,21].

Оценка точности определения глубины дефекта КРН приборами НК проводилась путем статистической обработки данных полученных при измерениях, согласно ГОСТ Р 8.736-2011. За оцениваемую величину, а именно значение, измеренное приборами НК глубины дефектов КРН, принимали среднее арифметическое из всех полученных отдельных измерений. Далее, для количественного выражения качества проведенного измерения каждым прибором НК, рассчитывали абсолютную и относительную погрешности определения глубины дефекта.

Результаты и их обсуждение

Сводная выборка данных по результатам измерений глубины всех представленных типов трещин КРН приборами НК приведена на рис. 1.

Линиями на графике ограничена область приемлемых отклонений результатов измерений глубины трещин КРН с относительной погрешностью измерений (δ) ± 30 % от действительного значения, определенного по результатам металлографического анализа. Количественное представление относительных погрешностей дефектов КРН приборами НК представлено на рис. 2. Из рис. 2 видно, что погрешность измерений приборов НК для ряда диапазонов глубин существенно различается, что объясняется наличием в контрольных точках измерений колоний трещин с разной длиной и плотностью, а также имеющих неправильный извилистый профиль.

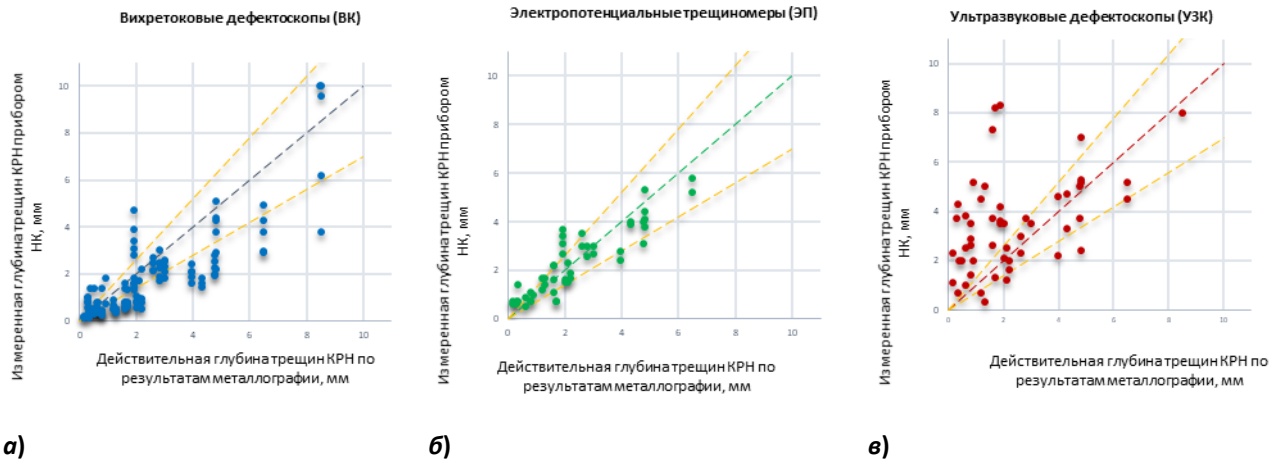


Рис. 1. Данные измерений глубины дефектов КРН приборами ВТД (а), ЭПТ (б), УЗД (в)

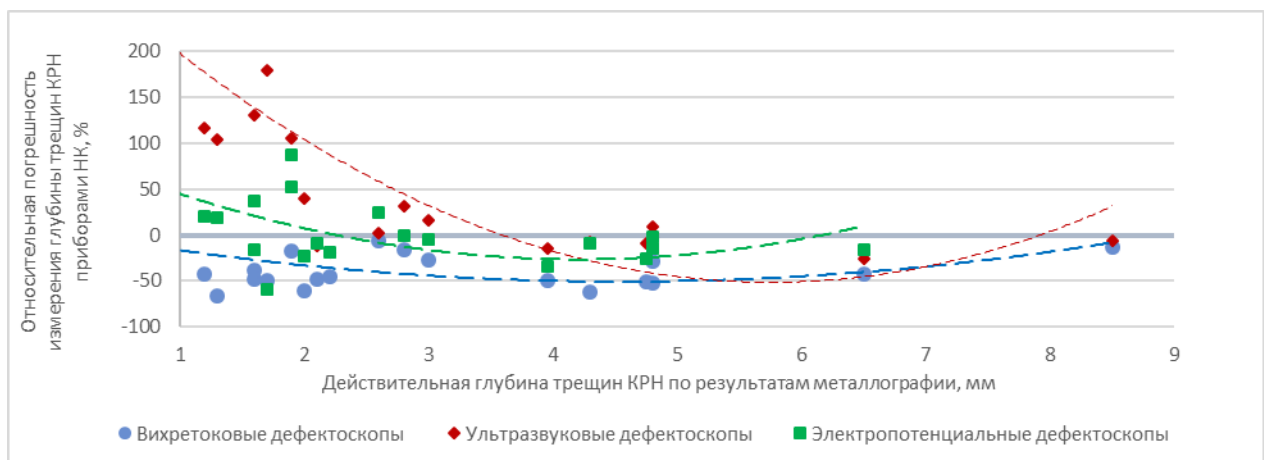


Рис. 2. Зависимости относительных погрешностей измерений дефектов КРН

Анализ представленных зависимостей показывает, что для представленных дефектов КРН малой глубины (до 1 мм) наиболее точные результаты показывают вихретоковые дефектоскопы. При использовании штатных вихретоковых дефектоскопов с ВТП $D \leq 3$ мм возможно получить приемлемую для практики погрешность оценки трещин КРН глубиной порядка до 3 мм. Приемлемые по точности результаты измерений в пределах $\delta \pm 30\%$, зафиксированы электро-потенциальными трещиномерами для глубин дефектов КРН более 1 мм и ультразвуковыми дефектоскопами при глубине дефектов более 2,5 мм.

На основании представленных результатов установлены зависимости влияния плотности колоний трещин КРН на достоверность оценки их глубины. При линейной аппроксимации значений прослеживается однозначная тенденция к возрастанию ошибки измерений при более плотном расположении трещин в колонии КРН (рис. 3).

Для ультразвуковых приборов определены наиболее перспективные для практики схемы контроля дефектов КРН: дифракции поперечных волн на трещине с использованием раздельно-совмещённого преобразователя и эхоимпульсной томографии с помощью ультразвукового дефектоскопа с цифровой фокусировкой апертуры антенной решётки.

С учетом удобства и практического применения приборов НК по выявлению и оценки параметров дефектов КРН в трассовых условиях являются предпочтительными вихретоковый и электропотенциальный методы. Вихретоковый контроль, как бесконтактный, целесообразно применять как поисковый, так и при первичной оценке глубины КРН в статическом режиме измерения. Применении электропотенциальных трещиномеров является наиболее приемлемым для определения глубин трещин КРН, но

более трудоемким и затратным при выполнении подготовительных операций по НК, так как потребуются дополнительная зачистка поверхности над дефектом до металлического блеска ($Rz \leq 40$ мкм по ГОСТ 2789-73), что существенно в трассовых условиях.

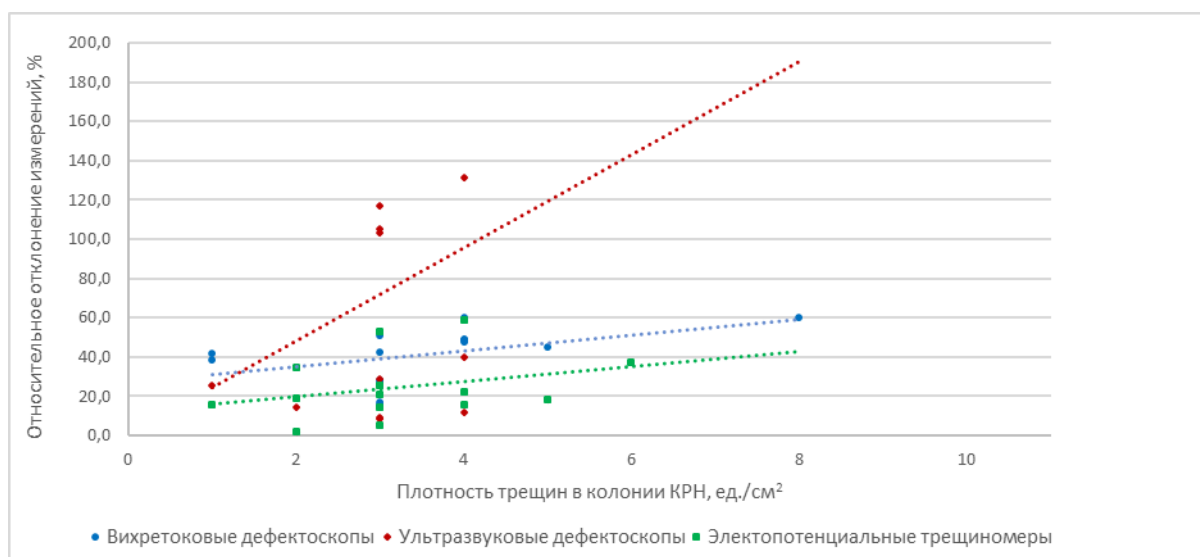


Рис. 3. Зависимость относительного отклонения измерений КРН исследуемыми методами НК от плотности трещин в колонии

При оценке глубины выявленной трещины большое значение имеет погрешность измерения в определенных диапазонах. По результатам сравнительного анализа измеренных значений глубины дефектов КРН приборами НК с результатами металлографических исследований, установлены погрешности измерений для ряда диапазонов фактических глубин трещин (таблица 1).

Таблица 1

Погрешности измерений приборов НК диапазонов фактических глубин трещин

Диапазон глубин трещин КРН, мм	Погрешности измерений					
	Вихретоковый метод НК (ВТП $D > 4$ мм)*		Вихретоковый метод НК (ВТП $D < 3$ мм)**		Электрический (электропотенциальный) метод НК	
	- Δ, %	+Δ, %	-Δ, %	+Δ, %	-Δ, %	+Δ, %
до 0,5 вкл.	- 66	184	- 41	65	-100	210
от 0,5 до 1,0	- 77	239	-37	63	-70	24
от 1,0 до 1,5	-63	60	- 48	38	-35	37
от 1,5 до 2,0	-60	65	-46	55	-41	41
от 2,0 до 2,5	-52	65	-48	17	-18	7
от 2,5 до 3,0	-54	62	-44	47	-35	37

* ВТП $D > 4$ мм – диаметр вихретокового преобразователя более 4 мм
 ** ВТП $D < 3$ мм – диаметр вихретокового преобразователя менее 3 мм

Практические рекомендации

На основании полученных результатов исследования разработаны практические рекомендации по оптимизации объемов НК в процессе категорирования и отбраковки труб:

- поиск дефектов КРН вихретоковым методом контроля, в динамическом режиме работы дефектоскопа с применением ВТП большого диаметра (ВТП $D \gg 4$ мм);
- локализация дефектной зоны с максимальными значениями вихретокового дефектоскопа, размером порядка 200×200 мм. Проведение магнитопорошковой дефектоскопии локализованного участка для визуализации и идентификации дефекта КРН;
- оценка глубины и ранжирование выявленных дефектов КРН вихретоковым методом контроля, с применением ВТП $D \leq 3$ мм в статическом режиме работы дефектоскопа;
- контрольное измерение глубины дефектов КРН, с применением электропотенциальных трещиномеров в местах максимальных показаний вихретоковых дефектоскопов при измеренной длине трещины более 50 мм и оценочной ее глубине более 1 мм;
- оценку глубины дефектов методом контролируемой шлифовки для дефектов КРН, находящимся в граничных интервалах с учетом установленных погрешностей приборов НК.

На рис. 4 показаны возможности применения данного подхода к задаче анализа труб магистрального газопровода $D_{н1420} \times 15,7$ мм с повреждениями КРН по глубине 10 % от толщины стенки труб с учетом приборных погрешностей, приведенных в таблице 1.

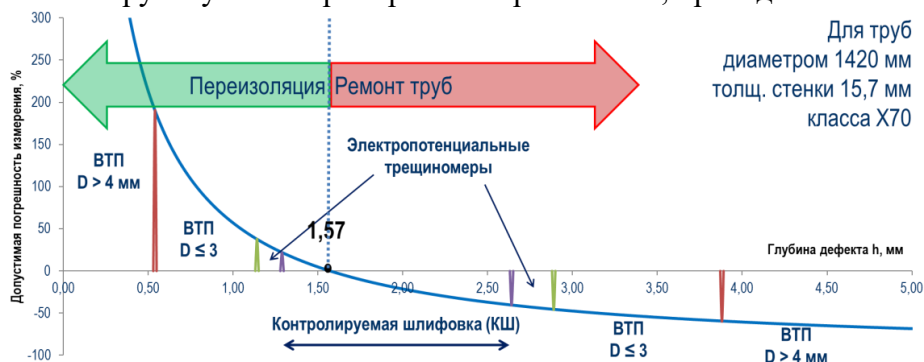


Рис. 4. Области применения приборов НК при отбраковке труб в процессе капремонта

Повреждения КРН, с оценочной глубиной и по показаниям вихретокового дефектоскопа с ВТП $D \leq 3$ мм до 1,1 мм относят к допустимым (менее 10% от толщины стенки трубы), а более 2,8 мм к недопустимым (более 10%). При этом статистическая доля дефектов КРН в данном диапазоне глубин в среднем составляет порядка 80 % от общего количества. Область применения электропотенциальных трещиномеров, как видно на рис.9, в данном интервале измерений достаточно мала, в связи с чем, применение данных приборов целесообразно только для контрольного подтверждения измерений вихретоковых приборов. Повреждения КРН, с оценочной глубиной от 1,1 мм до 2,8 мм включительно, по показаниям вихретокового дефектоскопа с ВТП $D \leq 3$ мм, относят к граничным. Статистическая доля дефектов КРН в данном диапазоне глубин в среднем составляет 18 % и оценку глубины для таких дефектов выполняют методом контролируемой шлифовки. Статистический анализ данных по повреждениям КРН выявленных в ходе неразрушающего контроля МГ ПАО «Газпром» в процессе их капитального ремонта, показывает, что глубина порядка 80% дефектов КРН от общего количества не превышает 10 % от толщины стенки трубы.

Выводы

На основании полученных результатов разработаны способы оценки размеров дефектов КРН и практические рекомендации по оптимизации объемов НК в процессе категорирования и отбраковки труб, позволяющие минимизировать затраты на контролирующую шлифовку труб при определении глубины дефектов КРН и сочетающие в себе высокую производительность технологических операций по дефектоскопии, достаточную чувствительность и достоверность оценки параметров дефектов КРН.

Библиографический список

1. Бычков Е.В., Крюков О.В. Исследование пульсаций потока газа в трубопроводах с использованием программного комплекса ANSYS CFX // В сборнике: Информационные системы и технологии ИСТ-2018. Материалы XXIV МНТК, посвященной 100-летию Нижегородской радиолоборатории. 2018. С. 654-660.
2. Васенин А.Б., Крюков О.В., Серебряков А.В. Энергоэффективные системы электроснабжения электроприводов нефтегазопроводов // В сборнике: Труды IX международной (XX Всероссийской) конференции АЭП-2016. Пермь. 2016. С. 380-384.
3. Крюков О.В. Комплексный анализ условий эксплуатации электродвигателей ГПА // Компрессорная техника и пневматика. 2013. № 4. С. 14-19.
4. Васенин А.Б., Крюков О.В. Проектирование электромеханической части и систем управления энергетических установок газотранспортных потребителей // Известия Тульского гос. ун-та. Технические науки. 2011. № 5-1. С. 47-51.
5. Васенин А.Б., Крюков О.В. Энергоэффективные и экологичные установки воздушного охлаждения // В сборнике: ВЕЛИКИЕ РЕКИ' 2017 труды научного конгресса 19-го МНПФ. НГАСУ. 2017. С. 93-96.
6. Крюков О.В., Васенин А.Б. Функциональные возможности энергетических установок при питании удаленных объектов // Электрооборудование: эксплуатация и ремонт. 2014. № 2. С. 50-56.
7. Крюков О.В., Серебряков А.В. Активно-адаптивные алгоритмы управления и мониторинга автономными энергетическими комплексами // В сборнике: Пром-Инжиниринг. Труды II МНТК. ЮУрГУ. 2016. С. 286-290.
8. Vasenin A.V., Kryukov O.V., Serebryakov A.V. Adaptive control algorithm soft autonomous generator complexes // В книге: Электромеханика, электротехнологии, электротехнические материалы и компоненты. Труды МКЭЭЭ-2016. 2016. С. 133-135.
9. Крюков О.В., Краснов Д.В. Перспективы применения преобразователей частоты для регулирования производительности ЭГПА // Газовая промышленность. 2014. № 6. С. 86-89.
10. Крюков О.В., Степанов С.Е., Серебряков А.В. Современный подход к организации ремонта по данным прогноза технического состояния и ресурса электрооборудования // Газовая промышленность. 2017. № 8 (756). С. 84-89.
11. Крюков О.В., Степанов С.Е. Модернизация систем управления ЭГПА в условиях действующих компрессорных станций // В сборнике: Проблемы автоматизации и управления в технических системах. МНТК под редакцией М.А. Щербакова. 2013. С. 29-32.
12. Крюков О.В. Мониторинг условий эксплуатации электродвигателей газоперекачивающих агрегатов // Контроль. Диагностика. 2016. № 12. С. 50-58.
13. Крюков О.В., Мещеряков В.Н., Туганов Р.Б. Экспериментальные исследования изменения магнитного поля трубопровода в зонах поверхностных дефектов // Автоматизация и ИТ в нефтегазовой области. 2020. № 2 (40). С. 24-29.

- 14.Крюков О.В. Методы мониторинга технического состояния электрооборудования КС // В сборнике: Наука. Технология. Производство-2016: Современные методы и средства диагностики электроэнергетического и электротехнического оборудования, средств и систем автоматики. Материалы ВНТК, посвященной 60-летию УГНТУ в Салавате. 2016. С. 149-153.
15. Васенин А.Б., Крюков О.В. Этапы организации поверки СИ расхода природного газа в рабочих условиях // В сборнике: Информационные системы и технологии ИСТ-2018. Материалы XXIV МНТК к 100-летию Нижегородской радиолaborатории. 2018. С. 668-674.
16. Васенин А.Б., Крюков О.В. Научные и технологические аспекты создания отраслевого метрологического расходоизмерительного центра ПАО "Газпром" // В сборнике: Информационные системы и технологии ИСТ-2018. Материалы XXIV МНТК, посвященной 100-летию Нижегородской радиолaborатории. 2018. С. 675-680.
17. Васенин А.Б., Степанов С.Е., Крюков О.В. Методология и средства оперативного мониторинга электродвигателей на КС // Контроль. Диагностика. 2019. № 11. С. 52-58.
18. Мещеряков В.Н., Ласточкин Д.В., Крюков О.В. Приложения теории нечетких множеств для обработки данных и задач прогнозирования в системах АЭП // В сборнике: Современные сложные системы управления. Материалы XII МНПК. 2017. С. 153-158.
- 19.Крюков О.В. Алгоритмы быстрого преобразования Уолша в микропроцессорных системах управления электроприводом // Известия ВУЗов. Электромеханика. 2005. № 4. С. 39-44.
- 20.Васенин А.Б., Степанов С.Е., Крюков О.В. Сравнительная оценка методов прогнозирования технического состояния электроприводов опасных производственных объектов// Контроль. Диагностика. 2020. Т. 23. № 11 (269). С. 54-62.
- 21.Захаров П.А., Туганов Р.Б., Крюков О.В. // Особенности проведения диагностики действующих трубопроводов в местах пересечения со строящимися газопроводами // Автоматизация и ИТ в нефтегазовой области. 2020. № 4 (42). С. 46-50.

S.E. Stepanov, A.B. Vasenin, A.S. Khlynin

METHODS OF "ON-LINE" PIPELINE CONDITION FORECASTING

Gazprom design LLC, Nizhny Novgorod

The results of research tests of nondestructive testing devices are presented in relation to the problem of estimating the size of colonies of stress-corrosion cracks on the surface of pipes that were in operation as part of main gas pipelines that differ in type size, properties and chemical composition. The regularities between the readings of devices that implement ultrasonic, eddy current, and electro potential control methods and the actual size of cracks are established. The reliability of the research results and conclusions is confirmed by the use of standardized methods of optical electron microscopy and X-ray tomography for verifying the size of cracks in the places of maximum readings of the devices.

Keywords: main gas pipeline, pipe surface, methods and devices for crack control, eddy current and electro potential monitoring methods.

ПРОГНОЗИРОВАНИЕ ТЕХНИЧЕСКОГО СОСТОЯНИЯ ТРАНСФОРМАТОРОВ РАСПРЕДЕЛИТЕЛЬНЫХ УСТРОЙСТВ

ООО «ТСН-электро», г. Нижний Новгород

Рассмотрены принципы формирования достоверных и надежных систем технической диагностики и прогнозирования трансформаторов распределительных устройств подстанций. Представлены теоретически обоснованные процедуры синтеза алгоритмов прогнозирования технического состояния распределительных устройств. Приведены примеры реализации методологии превентивного прогноза параметров трансформатора на одной из цифровых подстанций.

Ключевые слова: техническая диагностика, оперативное прогнозирование, техническое состояние, трансформатор, распределительное устройство, обслуживание по фактическому состоянию, алгоритм, диагностическая экспертная система.

Введение

В настоящее время все электроэнергетические системы (ЭЭС) с трансформаторными подстанциями (ТП) на различных объектах ТЭК относятся к промышленным объектам повышенной опасности [1-4] и независимо от режимов работы комплектных распределительных устройств (КРУ) с каждым годом федеральными и отраслевыми нормативно-техническими документами ужесточаются требования обеспечения повышенной надежности и безаварийности работы всего оборудования [5-8]. Это обусловлено не только объективными причинами (устаревшее оборудование, ненадежное электроснабжение и т.п.), но и нередко наличием неквалифицированного обслуживания и попыток внешнего проникновения. Особенно это актуально в связи с внедрением цифровых и интеллектуальных подстанций (рис. 1) [9-12].



Рис. 1. Комплектная трансформаторная подстанция в утепленной оболочке из сэндвич-панелей

В системах оперативного мониторинга технического состояния (ТС) ответственных агрегатов ЭЭС наиболее эффективной процедурой является не традиционное диагностирование трансформаторов КРУ ТП, а достоверное прогнозирование их ТС с оценкой остаточного ресурса агрегата [13-16]. Это позволяет предотвратить внезапные нештатные аварии и перейти от планово-предупредительных процедур технического обслуживания и ремонта (ТОиР) к обслуживанию по фактическому состоянию оборудования [17-20]. Методологический подход здесь может быть единым для любого электротехнического, электромеханического и электроэнергетического оборудования, технологических установок и систем [21-24].

Общий принцип управления техническим состоянием на основе прогнозирования состояния ТП состоит в том, что данные о состоянии объекта снимаются с датчиков и подаются на подсистему прогнозирования технического состояния.

В случае если формируется решение о наступающем отказе, то подсистема ТОиР выполняет действия, которые направлены на:

- предотвращение отказов, включая комплекс мероприятий по обеспечению функционирования системы без прерывания;
- минимизацию последствий отказа за счёт подготовки к ожидаемому отказу, что позволяет сократить время ремонта, а, следовательно, и продолжительность неработоспособного состояния.

Известно множество методов, которые основываются на определении объективных закономерностей развития дефектов и повреждений, статистической обработке данных, экстраполяции трендов до предельно допустимых значений и на вероятностной оценке значений показателей.

Методология систем прогнозирования технического состояния трансформаторов КРУ

В подсистеме прогнозирования определяется вероятность наступления отказа на рассматриваемом интервале времени. В условиях априорной неопределенности в решающем правиле используется оценка, которая находится в результате структурно-параметрического синтеза (обучения) предиктора на основе эмпирических данных, составляющих обучающую выборку. Решение о формировании сигнала-предупреждения о наступающем отказе выносится в результате сравнения с пороговым значением найденной вероятности наступления отказа на заданном интервале. Определение этого значения представляет отдельную задачу и связано с используемым критерием, который определяет эффективность прогнозов.

Применение прогнозирующего контроля направлено на сокращение издержек, вызванных простоем технического объекта. Поэтому эффект от внедрения прогнозирующего контроля можно оценить с помощью коэффициента технического использования. Эффект от применения прогнозирующего контроля охарактеризуем величиной, которая показывает, во сколько раз изменяется среднее время неработоспособного состояния технического объекта. Эффект от применения прогнозирующего управления техническим состоянием КРУ возрастает по мере уменьшения неотрицательной величины. Для проведения анализа эффективности применения процедур прогнозирования в составе системы управления техническим состоянием, охарактеризуем подсистему ТОиР вероятностью возникновения (не предотвращения) отказа при условии правильного предсказания и вероятностью привнесения отказа, при условии ошибочного предсказания (ложной тревоги).

Подсистема прогнозирования отказов может быть охарактеризована вероятностью ложной тревоги (ошибки первого рода) и вероятностью пропуска наступающего неисправного состояния технического объекта (ошибки второго рода). Если выполненные превентивные действия при правильном предсказании не позволили предотвратить отказ, то, как правило, устранение отказа выполняется за меньшее время.

Эффективность прогнозирования технического состояния возрастает для идеальной подсистемы ТОиР. При этом все правильно предсказанные отказы предотвращаются, а обработка ложных тревог не приводит к дополнительному простоям. В этом случае величина выигрыша полностью определяется вероятностью пропуска при прогнозировании отказа.

Достоинство применения прогнозирования технического состояния КРУ возрастает при высокой точности прогнозирования. Целевая функция представляет собой взвешенную сумму вероятностей ошибок первого и второго рода и по своей структуре близка к выражению для среднего риска.

Эффект от перехода к прогнозированию технического состояния КРУ

Выигрыш от применения прогнозирования уменьшается, при уменьшении априорной вероятности отказа на анализируемом интервале времени. При этом повышаются требования к точности прогнозирования. В то же время при использовании оптимального решающего правила не происходит снижения эффективности по отношению к системе ТОиР без прогнозирования технического состояния и определяется датчиками регистрации состояния цифровой ТП (рис. 3).

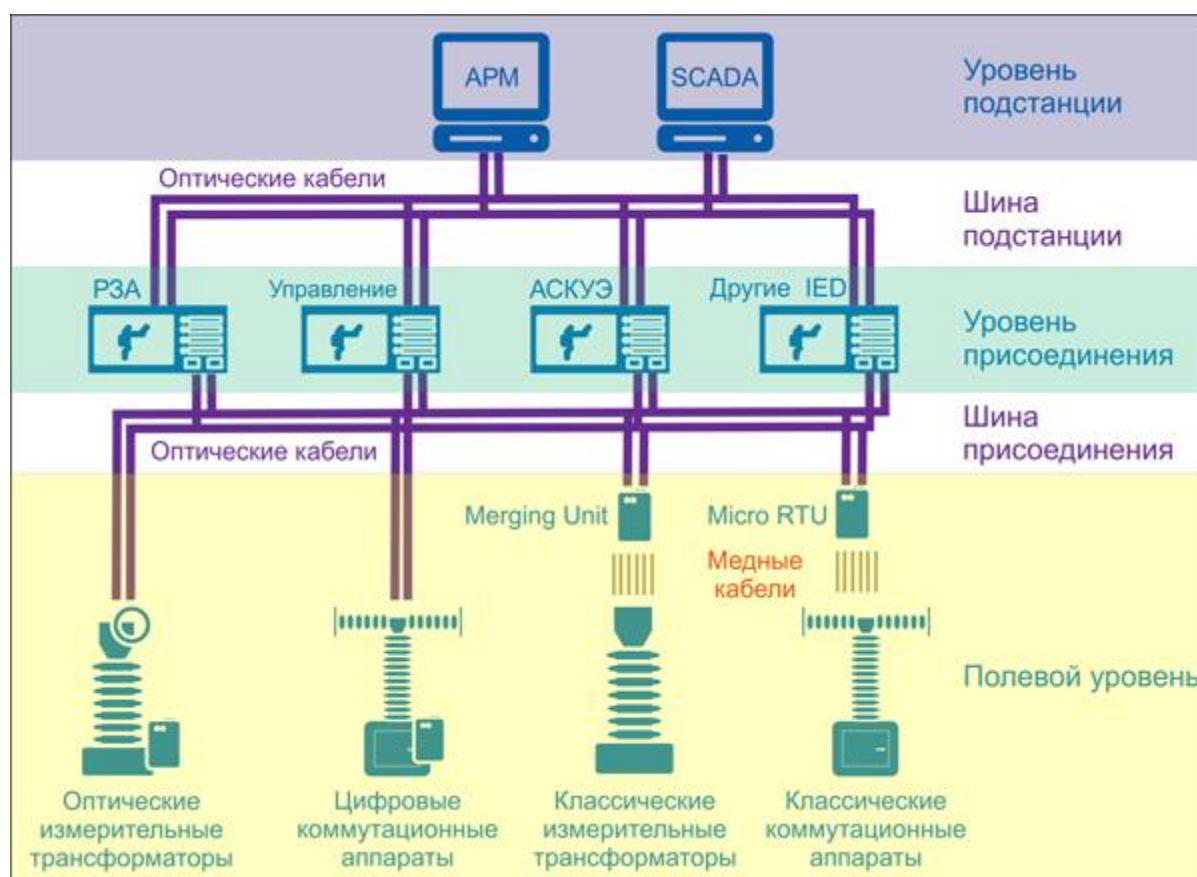


Рис. 3. Датчики регистрации состояния цифровой ТП

В общем случае информация о состоянии технического объекта может быть получена на основе обработки показаний интеллектуальных датчиков или результатов тестов. В случае использования вероятностной модели сложной системы элементарное действие по сбору информации заключается в определении значения одной из переменных. Такие действия характеризуются определенной стоимостью, представляющей затраты на проведение теста и ценностью. Однако на момент вынесения решения о целесообразности проведения некоторого теста его результат является неизвестным. Поэтому ожидаемый эффект от проведения теста определяется с использованием усреднения по множеству возможных исходов.

Проведение теста позволяет получить дополнительную информацию о состоянии технического объекта и, следовательно, приводит к меньшим значениям по сравнению с величиной, обеспечиваемой при использовании показаний датчиков. Сопоставляя сокращение издержек, вызванных простоем оборудования и стоимость теста можно определить ценность теста. Проведение дополнительных тестов завершается, если для любого из тестов их стоимость превышает ценность получаемой информации.

Алгоритм системы принятия решений по прогнозированию ТС ТП

Процедуру принятия решений о проведении процедур ТОиР представим с помощью сети принятия решений, которая содержит три типа вершин, и может рассматриваться как расширение байесовской сети.

Для упрощения используют сеть принятия решений, соответствующую единственному тесту. Вершины жеребьевки, обозначенные овалами, как и в байесовских сетях, представляют переменные (случайные величины). Вершины принятия решений (прямоугольники) представляют возможность лицу, принимающему решение выбрать одно из доступных действий на основе рассчитываемых значений целевой функции. В сетях принятия решений, основанных на критерии минимума среднего риска, вершины, обозначенные ромбами, содержат значения элементов матрицы потерь. В рамках развиваемого подхода последствия ошибочных решений и, соответственно, значения целевой функций зависят от характеристик подсистемы ТОиР.

В качестве примера возможного применения развиваемого подхода для управления техническим состоянием рассмотрим систему трансформаторной подстанции, информация о состоянии силового трансформатора которой регистрируется датчиками (рис. 4).

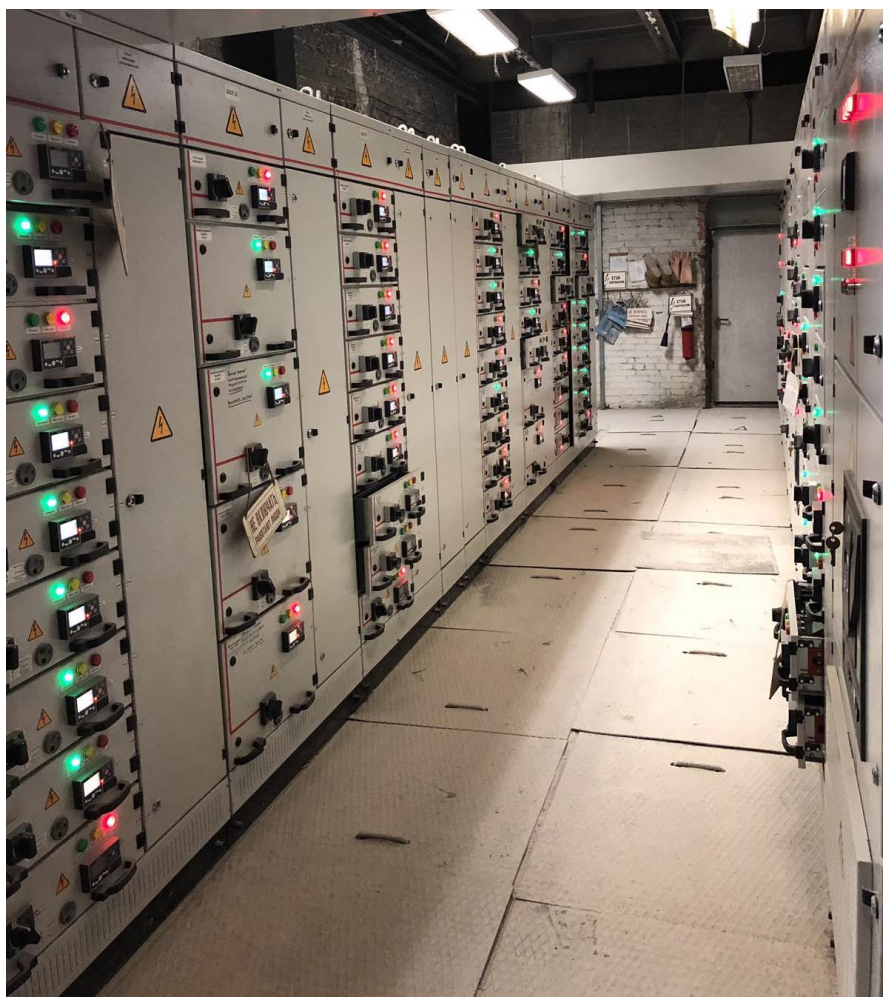


Рис. 4. Мониторинг КРУ ТП на оборудовании «Каскад»

Выделим три состояния такой технической системы: исправное, предотказное и неработоспособное. Если выносится решение, что система исправна, то действия по управлению состоянием системы не предпринимаются и функционирование системы продолжается.

В предотказном состоянии для предотвращения перехода системы в неработоспособное состояние выполняется автоматическое отключение с последующим восстановлением исправного состояния. При наличии существенной неопределенности относительно состояния системы может быть вынесено решение о проведении теста, результаты которого позволяют существенно повысить достоверность определения состояния системы. Реализация данных функций возможна на щитовом электрооборудовании «КРУ – Каскад» (рис. 4).

Выводы

Достоверное прогнозирование технического состояния особо ответственных технологических агрегатов с оценкой остаточного ресурса позволяет заблаговременно предотвратить внезапные нештатные аварии на объектах повышенной опасности и перейти от планово-предупредительных ТОиР к обслуживанию по фактическому состоянию оборудования со значительным экономическим эффектом.

В общем случае информация о техническом состоянии трансформаторных агрегатов КРУ ТП получается на основе обработки показаний интеллектуальных датчиков и результатов тестов. При этом выигрыш от применения процедур прогнозирования увеличивается, при росте априорной вероятности отказов на анализируемом интервале времени и повышении требований к точности прогнозирования.

Библиографический список

1. Репин Д.Г., Крюков О.В. Концепты системы мониторинга технического состояния компрессорных станций // Контроль. Диагностика. 2017. № 12. С. 30-35.
2. Крюков О.В., Серебряков А.В., Макриденко Л.А., Волков С.Н., Сарычев А.П., Кобельков Н.О. Мониторинг и прогнозирование технического состояния электромеханических систем энергетики // М.: АО «Корпорация ВНИИЭМ». – 2017.
3. Васенин А.Б., Крюков О.В. Энергоэффективные и экологичные установки воздушного охлаждения // В сборнике: ВЕЛИКИЕ РЕКИ' 2017 труды научного конгресса 19-го МНПФ. НГАСУ. 2017. С. 93-96.
4. Васенин А.Б., Крюков О.В. Проектирование электромеханической части и систем управления энергетических установок газотранспортных потребителей // Известия Тульского гос. ун-та. Технические науки. 2011. № 5-1. С. 47-51.
5. Крюков О.В. Мониторинг условий эксплуатации электродвигателей газоперекачивающих агрегатов // Контроль. Диагностика. 2016. № 12. С. 50-58.
6. Крюков О.В. Комплексный анализ условий эксплуатации электродвигателей ГПА // Компрессорная техника и пневматика. 2013. № 4. С. 14-19.
7. Крюков О.В., Степанов С.Е., Васенин А.Б. Поддержка диспетчерских решений ГТС на базе оценки их энергоэффективности // Наука и техника в газовой промышленности. 2019. № 4 (80). С. 71-81.
8. Вожаков А.В., Крюков О.В., Лисин Н.Г. О причинах недостаточно эффективного построения и эксплуатации систем автоматизации // Автоматизация в промышленности. 2012. № 2. С. 38-47.
9. Крюков О.В. Система оперативно-диспетчерского управления подстанциями цехов // Электрооборудование: эксплуатация и ремонт. 2014. № 1. С. 43-45.
10. Кононенко А.Б., Косоротов А.А., Крюков О.В. Расширение функциональных возможностей автоматизации и мониторинга распределительных устройств КТП «Каскад» // Автоматизация и ИТ в энергетике. 2020. № 12. С. 26-31.
11. Реализованные проекты ООО «ТСН-электро». Режим доступа: <https://www.tcn-nn.ru/>

12. Kryukov O.V., Blagodarov D.A., Dulnev N.N., et al. Intelligent control of electric machine drive systems // В сборнике: 2018 10th International Conference on Electrical Power Drive Systems, ICEPDS 2018 - Conference Proceedings 10. 2018. С. 8571670.
13. Степанов С.Е., Крюков О.В. Выбор методов мониторинга и прогнозирования технического состояния автоматизированных электроприводов энергетических объектов // Контроль. Диагностика. 2018. № 11. С. 32-39.
14. Крюков О.В. Алгоритмы быстрого преобразования Уолша в микропроцессорных системах управления электроприводом // Известия ВУЗов. Электромеханика. 2005. № 4. С. 39-44.
15. Vasenin A.V., Kryukov O.V., Serebryakov A.V. Adaptive control algorithms of autonomous generator complexes // В книге: Электромеханика, электротехнологии, электротехнические материалы и компоненты. Труды МКЭЭЭ-2016. 2016. С. 133-135.
16. Крюков О.В., Серебряков А.В. Активно-адаптивные алгоритмы управления и мониторинга автономными энергетическими комплексами // В сборнике: Пром-Инжиниринг. Труды II МНТК. ЮУрГУ. 2016. С. 286-290.
17. Крюков О.В., Степанов С.Е., Серебряков А.В. Современный подход к организации ремонта по данным прогноза технического состояния и ресурса электрооборудования // Газовая промышленность. 2017. № 8 (756). С. 84-89.
18. Васенин А.Б., Степанов С.Е., Крюков О.В. Методология и средства оперативного мониторинга электродвигателей на КС // Контроль. Диагностика. 2019. № 11. С. 52-58.
19. Бычков Е.В., Крюков О.В. Теоретические и аппаратные возможности перехода на обслуживание по фактическому состоянию // В сб.: Информационные системы и технологии ИСТ-2018. Мат. XXIV МНТК к 100-летию Нижегородской радиолоборатории. 2018. С. 661-667.
20. Крюков О.В. Прогнозирование теплового режима двигателей СТД-12500 методом спектрального анализа // В сборнике: Актуальные проблемы электроэнергетики XXXII региональной НТК, посвященной 75-летию проф. С.В. Хватова. Н. Новгород, 2013. С. 140-146.
21. Крюков О.В., Степанов С.Е. Модернизация систем управления ЭГПА в условиях действующих компрессорных станций // В сборнике: Проблемы автоматизации и управления в технических системах. МНТК под редакцией М.А. Щербакова. 2013. С. 29-32.
22. Мещеряков В.Н., Ласточкин Д.В., Крюков О.В. Приложения теории нечетких множеств для обработки данных и задач прогнозирования в системах АЭП // В сборнике: Современные сложные системы управления. Материалы XII МНПК. 2017. С. 153-158.
23. Крюков О.В., Мещеряков В.Н., Гуляев И.В. Электроприводы на основе машины двойного питания и асинхронного вентильного каскада с преобразователями в цепях статора и ротора // Саранск, 2020.
24. Kryukov O.V., Gulyaev I.V., Terlukhov D.Y. Method for stabilizing the operation of synchronous machines using a virtual load sensor // Russian Electrical Engineering. 2019. T. 90. № 7. С. 473-478.

A.S. Volkov, V.V. Gruzdev, O.V. Kryukov

FORECASTING OF THE TECHNICAL CONDITION OF TRANSFORMERS OF SWITCHGEARS

TSN-electro LLC, Nizhny Novgorod

The principles of forming reliable and reliable systems of technical diagnostics and forecasting of substation switchgear transformers are considered. Theoretically justified procedures for the synthesis of algorithms for predicting the technical condition of switchgear are presented. Examples of the implementation of the methodology of preventive prediction of transformer parameters at one of the digital substations are given.

Keywords: technical diagnostics, operational forecasting, technical condition, transformer, switchgear, maintenance according to the actual condition, algorithm, diagnostic expert system.

МОНИТОРИНГ КТП «КАСКАД» ДЛЯ ОБСЛУЖИВАНИЯ ПОДСТАНЦИЙ ПО ФАКТИЧЕСКОМУ СОСТОЯНИЮ

ООО «ТСН-электро», г. Нижний Новгород

Рассмотрены перспективы внедрения новых технологий технического обслуживания для расширения функциональных возможностей и надежности работы комплектных трансформаторных подстанций. Предложены инновационные функции мониторинга и интеллектуального управления автоматизированными системами автоматического ввода резерва распределительных устройств в рамках концепции цифровых подстанций. Приведены примеры новых функций системы комплектных распределительных устройств НКУ «Каскад» на базе контроллеров НМИ с использованием веб-технологий.

Ключевые слова: цифровая подстанция, распределительное устройство, автоматический ввод резерва, техническое обслуживание, интеллектуализация.

Введение

В настоящее время на предприятиях, эксплуатирующих электротехнические шкафы комплектных распределительных устройств (КРУ), распространен способ планового обслуживания электрооборудования (ППР). При этом вне зависимости от реального состояния отдельных элементов и систем электрооборудования оно в установленном плановом порядке выводится из эксплуатации, снижая технико-экономические показатели. Сегодня подобная технология технического обслуживания и ремонта (ТОиР) электроэнергетических систем является затратной и не обеспечивает адекватную достоверную оценку состояния и остаточного ресурса электрооборудования [1-3]. Для повышения надежности, безопасности и энергоэффективности эксплуатации электротехнического оборудования КРУ необходима комплексная система оперативного мониторинга и прогнозирования технического состояния, интегрированная в АСУ ТП цифровой подстанции (ЦТП) в режиме “on-line” [4-6]. Подобная технология ТОиР по фактическому состоянию электрооборудования позволяет за счет постоянного контроля основных (критических) параметров и обработки их в соответствии с интеллектуальными алгоритмами прогнозирования остаточного ресурса выдавать эксплуатационному персоналу превентивные данные о нарастании рисков неисправностей.

Для перехода от ППР электрооборудования КРУ ЦТП к ТОиР по фактическому состоянию требуется не только достоверное отслеживание текущего состояния всего комплекса коммутационных и измерительных аппаратов, установленных внутри шкафов КРУ, но и оперативно информировать операторов о необходимости проведения регламентных профилактических работ конкретных узлов. Это обуславливает поиск инновационных технологий на основе интеллектуализации всех составляющих, включая современные низковольтные (0,4 и 0,69 кВ) КРУ ЦТП [7-9]. Активное внедрение в энергетику сетевых технологий и увеличение доступности средств автоматизации связано с применением программируемых контроллеров и человеко-машинного интерфейса (ЧМИ) или *Human-MachineInterface, НМИ*.

Основным компонентом автоматизации КРУ является система автоматического ввода резерва (АВР). Главным интеллектуальным центром ее является контроллер АВР, который представляет собой специализированный блок (например, БМРЗ-0,4 [10]) или систему свободно программируемых устройств (например, ПР [11-12] или ПЛК [13-14]).

Уровни интеллектуализации КРУ

Применение специализированных блоков АВР таких как БМРЗ-0,4, OptiSave или LovatoATL сопряжено с определенными трудностями:

- в случае с блоками БМРЗ-0,4 - это высокие стоимость и сложность интеграции;
- в случае с более простыми блоками - отсутствие достаточно гибкой настройки конкретные задачи, трудности в случае расширения функциональных возможностей вследствие невозможности вносить изменения в алгоритмы работы данных устройств. Кроме того, при повреждении входа или выхода устройства необходимо менять его целиком.

Использование программируемых реле (ПР) позволяет реализовывать логику работы под каждую конкретную задачу, причем достаточно простыми программными средствами. Также при наличии резерва, можно перепрограммировать входы/выходы ПР в случае выхода их из строя. Практически все подобные ПР имеют на борту коммуникационные интерфейсы (SiemensLogo, ОВЕН ПР200) или коммуникационные модули расширения (ZelioLogic) для передачи данных в системы диспетчеризации. Но ПР подходят лишь для малой мощности.

Применение современных промышленных контроллеров среднего уровня автоматизации (серия ПЛК210 «ОВЕН», M241, M251, M262 «SchneiderElectric», SiemensS7-1200, WAGO PFC200) позволяет существенно расширить возможности системы АВР РУ. Обладая достаточно серьезной производительностью, данные контроллеры берут на себя помимо задач непосредственно АВР еще и задачи организации телеизмерения и телеуправления, фиксации аварийных событий, а также взаимодействие с пользователем посредством НМИ. Именно такая перспективная идеология применена в системе автоматизированного мониторинга и диагностики КРУ «Каскад» от компании «ТСН-электро» [15].

Описание системы мониторинга и управления «КРУ-Каскад»

Конструктивно КРУ представляет собой программно-аппаратный комплекс (рис. 1), обеспечивающий мониторинг основных технологических параметров ЦТП, управление аппаратами главных цепей и предоставление рекомендаций о необходимости ТОиР для оборудования в КРУ.

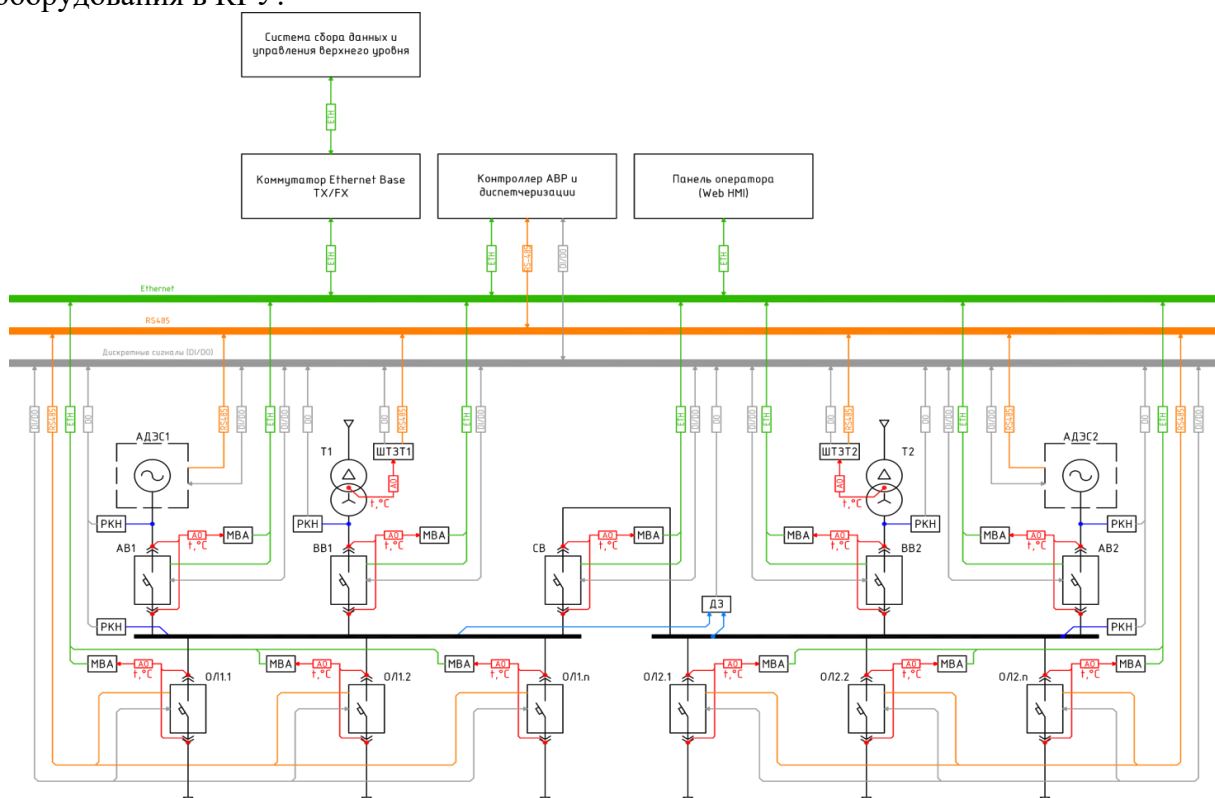


Рис. 1. Структурная схема системы диспетчеризации НКУ РУ «Каскад»

Система устанавливается в каждый шкаф КРУ с вакуумным выключателем и управлением от отечественного ПЛК с сенсорным цветным графическим экраном, обеспечивая непрерывную проверку технического состояния оборудования шкафа. Совокупные функциональные возможности КРУ «Каскад» обеспечивают адекватное оповещение персонала о необходимости профилактического ТОиР оборудования. Данное КРУ представляет собой две секции шин, к которым подключены автоматические выключатели (АВ) отходящих линий (ОЛ). Секции питаются либо от ввода через силовой трансформатор (Т1, Т2), либо от автоматизированных дизельных электрических станций (АДЭС1, АДЭС2). Защиту вводов осуществляют АВ типа ВВ1 и ВВ2 (для вводов от трансформаторов) и АВ1, АВ2 (для вводов от АДЭС). Секционирование происходит выключателем СВ. Контроль нормального напряжения осуществляют реле контроля напряжения (РКН), которые расположены на вводах и секции шин. Дополнительно на секциях шин предусмотрены устройства дуговой защиты для оперативного отключения питания на секции шин при возникновении дуги. Защиту трансформаторов осуществляют устройства тепловой защиты ШТЗТ1 и ШТЗТ2. Для осуществления контроля температуры контактных соединений автоматических выключателей применяются датчики температуры, которые подключены к аналоговым модулям ввода (МВА).

Данные о состоянии электрооборудования РУ поступают в контроллер АВР и диспетчеризации [16-18]. Для этого используются различные каналы передачи информации:

1. Дискретные сигналы состояния (состояние коммутационных аппаратов, РКН, блока ДЗ, аварийная температура трансформатора, состояние АДЭС и т.д.) и управления (включение, отключение автоматических выключателей, пуск и останов АДЭС).
2. Интерфейс RS-485 (температура обмоток трансформатора, измерения с электронных расцепителей аппаратов отходящих линий, дополнительная информация от АДЭС).
3. Интерфейс Ethernet (данные с электронных расцепителей вводных и секционного автоматов, температура контактных соединений от блоков аналогового ввода) [19].

Для взаимодействия с пользователем (визуализация и изменение параметров работы системы) используется панель оператора, которая подключена к контроллеру посредством Ethernet. Для связи с системой верхнего уровня может быть установлен коммутатор (маршрутизатор) с поддержкой интерфейса Ethernet, BaseTX, так и BaseFX (оптический канал).

Функциональные особенности системы.

Главными приоритетными преимуществами новой системы НКУ «Каскад» является комплекс инновационных функций, которые дают синергетический эффект от применения КТП:

Визуализация данных на НМІ при помощи веб-технологий. Состояние электрооборудования наглядно отображается на мнемосхеме (рис. 2).

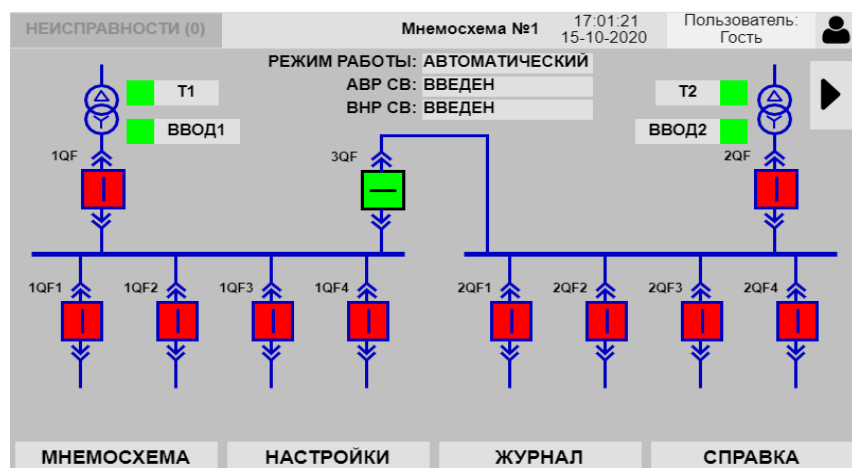


Рис. 2. Мнемосхема РУ-0,4 кВ

При этом можно отслеживать состояние и положение коммутационных аппаратов распределительного устройства, состояние трансформаторов (температура и срабатывание защиты), наличие нормального напряжения на вводах и секциях шин, а также наличие и количество текущих неисправностей. НМИ разработан с применением веб-интерфейса и доступен как с панели оператора на самом распределительном устройстве, так и может быть отображен удаленно на ПК с любым веб-браузером с поддержкой HTML5, что позволяет проводить работы с контроллером даже при отключенной или же неисправной панели оператора.

Реализация различных алгоритмов системы автоматического ввода резерва под существующие схемы энергоснабжения. Опираясь на многолетний опыт изготовления и внедрения КТП с системами АВР авторами проработаны алгоритмы, которые учитывают различные ситуации, схемы аварийного ввода резерва и автоматического возврата нормального режима. Гибкие настройки позволяют адаптировать ПО контроллера АВР посредством установки режимов подключения, отключения и переключения основных и резервных вводов, режимов пуска и останова АДЭС, а также определения реакции блока АВР на выявление неисправностей (рис. 3). Все это позволяет с минимальными временными затратами корректировать алгоритм работы под конкретные нужды эксплуатационного персонала.



Рис. 3. Настройки АВР

Контроль температуры контактных соединений. Система диагностики КТП «Каскад» может комплектоваться датчиками температуры, которые устанавливаются на места контактных соединений коммутационных аппаратов, с целью диагностики их состояния. Датчики подключаются к контроллеру, который оценивает абсолютное значение и изменения температуры, и в случае превышения критических параметров, рассчитывает время до отключения коммутационного аппарата. Пороги уровней температуры задаются в настройках.

Журналы текущих неисправностей и журнал событий. Контроллер системы диагностики отслеживает и записывает во внутреннюю энергонезависимую память все события, связанные с изменением состояния электрооборудования и настроек системы (рис. 4, а и б).

НЕИСПРАВНОСТИ (3)		РУНН 0,4кВ	Неисправности	17:27:57 15-10-2020	Пользователь: Гость
Дата и время	Подтверждение	Неисправность			
1	15-10-2020 17:24:09.003	1QF.Сработал расцепитель			
2	15-10-2020 17:24:03.008	2QF.Отключен АВ опер. цепей 24 VDC			
3	15-10-2020 17:23:52.007	1QF.Отключен АВ опер. цепей 24 VDC			

КВИТИРОВАНИЕ

МНМОСХЕМА НАСТРОЙКИ ЖУРНАЛ СПРАВКА

НЕИСПРАВНОСТИ (0)		Журнал событий	17:12:23 15-10-2020	Пользователь: Гость
Дата и время	Событие			
1	15-10-2020 16:58:34.057	СШ1.Напряжение восстановлено		
2	15-10-2020 16:58:34.057	1QF.Включен		
3	15-10-2020 16:58:34.031	СШ1.Напряжение отсутствует		
4	15-10-2020 16:58:34.031	3QF.Отключен		
5	15-10-2020 16:58:29.017	ВВ1.Напряжение восстановлено		
6	15-10-2020 16:58:06.009	СШ1.Напряжение восстановлено		
7	15-10-2020 16:58:06.009	3QF.Включен		
8	15-10-2020 16:58:06.009	Сработал АВР СВ		
9	15-10-2020 16:58:06.004	1QF.Отключен		
10	15-10-2020 16:58:03.008	СШ1.Напряжение отсутствует		
11	15-10-2020 16:58:03.008	ВВ1.Напряжение отсутствует		
12	15-10-2020 16:55:15.081	АВР СВ введен		

МНМОСХЕМА НАСТРОЙКИ ЖУРНАЛ СПРАВКА

а) б)
Рис. 4. Список (журнал) неисправностей (событий)

Каждому событию присваивается метка времени с точной привязанной ко времени скана контроллера. В отличие от журналов, основанных на базе стандартных НМИ, подобная реализация с высокой точностью отслеживает последовательность событий до аварии.

Самодиагностика алгоритма АВР и выявление неявных неисправностей. В некоторых случаях может произойти ситуация, при которой контроллер АВР не выявил явных неисправностей оборудования, однако алгоритм обрабатывает некорректно или не полностью («зависает» на полпути). Это может быть, если на контроллер не приходит сигнал, который обычно был. Поэтому реализована возможность просмотра каких же условий.

Переназначение входов и выходов контроллера. В случае выявления неисправности входа или выхода контроллера имеется возможность переназначения неисправного канала на неиспользуемый (резервный) вход или выход (рис. 5). Это позволяет оперативно исправлять подобные неисправности без перепрограммирования контроллера.

НЕИСПРАВНОСТИ (3)		РУНН 0,4кВ	Настройки	17:26:53 15-10-2020	Пользователь: Гость		
Вход	Уст.	(№ выхода по схеме)	Сигнал	Вход	Уст.	(№ выхода по схеме)	Сигнал
I1	<input type="checkbox"/>	1.0.1	(0.1)1QF.Включен	I13	<input type="checkbox"/>	1.0.13	(0.13)2QF.Включен
I2	<input checked="" type="checkbox"/>	1.0.2	(0.2)1QF.Отключен	I14	<input type="checkbox"/>	1.0.14	(0.14)2QF.Выкачен
I3	<input type="checkbox"/>	1.0.3	(0.3)1QF.Аварийно отключен	I15	<input type="checkbox"/>	1.0.15	(0.15)2QF.Готовность к включению
I4	<input checked="" type="checkbox"/>	1.0.4	(0.4)1QF.Включен	I16	<input checked="" type="checkbox"/>	1.0.16	(0.16)ВВ2.Напряжение на вводе исправно
I5	<input type="checkbox"/>	1.0.5	(0.5)1QF.Выкачен	I17	<input type="checkbox"/>	1.0.17	(0.17)СШ2.Напряжение на СШ исправно
I6	<input type="checkbox"/>	1.0.6	(0.6)1QF.Готовность к включению	I18	<input type="checkbox"/>	1.0.18	(0.18)Тр2.Аварийная температура
I7	<input type="checkbox"/>	1.0.7	(0.7)ВВ1.Напряжение на вводе исправно	I19	<input type="checkbox"/>	1.0.19	(0.19)3QF.Включен
I8	<input type="checkbox"/>	1.0.8	(0.8)СШ1.Напряжение на СШ исправно	I20	<input checked="" type="checkbox"/>	1.0.20	(0.20)3QF.Отключен
I9	<input type="checkbox"/>	1.0.9	(0.9)Тр1.Аварийная температура	I21	<input type="checkbox"/>	1.0.21	(0.21)3QF.Аварийно отключен
I10	<input checked="" type="checkbox"/>	1.0.10	(0.10)2QF.Включен	I22	<input checked="" type="checkbox"/>	1.0.22	(0.22)3QF.Включен
I11	<input type="checkbox"/>	1.0.11	(0.11)2QF.Отключен	I23	<input type="checkbox"/>	1.0.23	(0.23)3QF.Выкачен
I12	<input type="checkbox"/>	1.0.12	(0.12)2QF.Аварийно отключен	I24	<input checked="" type="checkbox"/>	1.0.24	(0.24)3QF.Готовность к включению

Дискретные входы ПЛК А1.0 (TM241CE40R) **НАЗНАЧЕНИЕ ВХОДОВ**

МНМОСХЕМА НАСТРОЙКИ ЖУРНАЛ СПРАВКА

Рис. 5. Переназначение входов и выходов

Разграничение прав пользователей. Для ограничения доступа к настройкам системы диагностики предусмотрена система прав пользователей по категориям.

Интеграция в системы верхнего уровня по различным протоколам связи. Система диагностики «Каскад» поддерживает передачу данных по протоколам связи ModBusRTU, ModBusTCP [20], МЭК 60870-5-104 [21], МЭК 60870-5-101 [22], МЭК 61850-8-1 (MMS) [23] в зависимости от выбранной платформы ПЛК. Также контроллер может служить концентратором данных и преобразователем протоколов, собирая данные от других устройств подстанции, а также от устройств тепловой защиты трансформаторов.

Применение стандартизированной среды разработки. Для разработки и отладки программного обеспечения системы диагностики «Каскад» используются ПО на базе системы программирования для ПЛК CODESYS 3.5 с использованием языков программирования, описанных стандартом МЭК 61131-3 [24] единой среды разработки.

Таким образом, основными преимуществами системы «КРУ-Каскад» являются:

- Наглядное представление информации о техническом состоянии аппаратов КРУ.
- Повышение надежности работы всего коммутационного оборудования и сокращение расходов эксплуатации за счет непрерывного контроля технического состояния.
- Переход от планового ТОиР к обслуживанию по фактическому состоянию.
- Доступ к документации по регламенту ТОиР и руководствам по эксплуатации.
- Интеграция в Единую систему диспетчеризации ЦТП по ModbusRTU/TCP.

Выводы

Внедрение системы диагностики «КРУ-Каскад» позволяет перейти к ТОиР по фактическому состоянию электрооборудования и продлить срок службы электрооборудования ЦТП за счет предупреждения аварийных ситуаций путем своевременной реакции на них. Кроме того, облегчается поиск причин неполадок (явных и скрытых) персоналом, упрощается наладка и ввод в строй. Широкие коммуникационные возможности позволяют интегрировать систему диагностики в комплекс управления производством предприятия.

Библиографический список

1. Стратегия развития электросетевого комплекса РФ: Утв. распоряжением Правительства РФ от 3 апреля 2013 г. № 511-р; в ред. постановления Правительства РФ от 29 ноября 2017 г. № 2664-р // Собр. законодательства РФ. 2017. № 49. С. 21027-21029.
2. Васенин А.Б., Степанов С.Е., Крюков О.В. Методология и средства оперативного мониторинга электродвигателей на КС // Контроль. Диагностика. 2019. № 11. С. 52-58.
3. Крюков О.В. Мониторинг условий эксплуатации электродвигателей газоперекачивающих агрегатов // Контроль. Диагностика. 2016. № 12. С. 50-58.
4. Программа инновационного развития ПАО «ФСК ЕЭС» на 2016-2020 годы с перспективой до 2025 года. [Электронный ресурс] М.: ФСК ЕЭС. 2016. Режим доступа: http://www.fsk-ees.ru/upload/docs/2_2016_PIR_FSK-2016-2020-2025.pdf
5. Крюков О.В., Серебряков А.В. Активно-адаптивные алгоритмы управления и мониторинга автономными энергетическими комплексами // В сборнике: Пром-Инжиниринг. Труды II МНТК. ЮУрГУ. 2016. С. 286-290.
6. Мещеряков В.Н., Ласточкин Д.В., Крюков О.В. Приложения теории нечетких множеств для обработки данных и задач прогнозирования в системах АЭП // В сборнике: Современные сложные системы управления. Материалы XII МНПК. 2017. С. 153-158.
7. Основные положения концепции интеллектуальной энергосистемы с активно-адаптивной сетью. [Электронный ресурс] М.: ФСК ЕЭС. 2012. Режим доступа: http://www.fsk-ees.ru/upload/docs/ies_aas.pdf.
8. Крюков О.В. Алгоритмы быстрого преобразования Уолша в микропроцессорных системах управления электроприводом // Известия ВУЗов. Электромеханика. 2005. № 4. С. 39-44.
9. Vasenin A.B., Kryukov O.V., Serebryakov A.V. Adaptive control algorithms of autonomous generator complexes // В книге: Electromechanics, Electrotechnology, Electromaterials and Components. Труды МКЭЭЭ-2016. М.: МЭИ (ТУ). 2016. С. 133-135.
10. Принципиальные электрические схемы защиты и автоматики присоединений 0,4 кВ для КТП-10(6)/0,4 на постоянном оперативном токе, выполненных с применением блоков серии БМРЗ-0,4. ТИ-031-2018. Электронный ресурс: https://www.mtrele.ru/files/project/albom_shem_zasch_i_avto_prisoed_04_pot_s_primeneniem_bmrz-04.pdf

11. Техническая коллекция SchneiderElectric. Выпуск № 18 Типовые схемы АВР с применением интеллектуального программируемого реле ZelioLogic. Электронный ресурс: <http://www.netkom.by/docs/N18-Tipovye-shemy-AVR-s-primeneniem-Zelio-Logic.pdf>
12. Щит автоматического ввода резерва АВР-3-11. Электронный ресурс: https://owen.ru/project/shhit-avtomaticheskogo-vvoda-rezerva-avr-3-11_
13. Типовое решение АВР. Электронный ресурс: <https://schneider-electric.app.box.com/s/p25u8feunpu42fflpoi80mlg54cof641/file/308512304640>
14. Васенин А.Б., Крюков О.В. Проектирование электромеханической части и систем управления энергетических установок // Известия ТГУ. Технические науки. 2011. № 5-1. С. 47-51.
15. Реализованные проекты ООО «ТСН-электро». Режим доступа: <https://www.tcn-nn.ru/>
16. Крюков О.В., Степанов С.Е. Модернизация систем управления ЭГПА в условиях действующих компрессорных станций // В сборнике: Проблемы автоматизации и управления в технических системах. МНТК под редакцией М.А. Щербакова. 2013. С. 29-32.
17. Васенин А.Б., Крюков О.В. Энергоэффективные и экологичные установки воздушного охлаждения // В сборнике: Великие реки' 2017 труды научного конгресса 19-го МНПФ. НГАСУ. 2017. С. 93-96.
18. Крюков О.В., Степанов С.Е., Серебряков А.В. Современный подход к организации ремонта по данным прогноза технического состояния и ресурса электрооборудования // Газовая промышленность. 2017. № 8 (756). С. 84-89.
19. Крюков О.В. Коммуникационная среда передачи данных сети ETHERNET на полевом уровне различных объектов // Автоматизация в промышленности. 2012. № 12. С. 26-30.
20. MODBUS APPLICATION PROTOCOL SPECIFICATION V1.1b3. Электронный ресурс: https://modbus.org/docs/Modbus_Application_Protocol_V1_1b3.pdf.
21. ГОСТ Р МЭК 60870-5-104-2004. http://docs.cntd.ru/document/gost-r-mek-60870-5-104-2004_
22. ГОСТ Р МЭК 60870-5-101-2006. http://docs.cntd.ru/document/gost-r-mek-60870-5-101-2006_
23. ГОСТ Р МЭК 61850-7-1-2009. Сети и системы связи на подстанциях. Часть 7. Базовая структура связи для подстанций и линейного оборудования. Раздел 1. Принципы и модели <http://docs.cntd.ru/document/1200082215>
24. ГОСТ Р МЭК 61131-3-2016 Контроллеры программируемые. Часть 3. Языки программирования <http://docs.cntd.ru/document/1200135008>.

A.B. Kononenko, A.A. Kosorotov, O.V. Kryukov

MONITORING OF KTP "CASCADE" FOR SUBSTATION MAINTENANCE ACCORDING TO THE ACTUAL CONDITION

TSN-electro LLC, Nizhny Novgorod

The prospects for the introduction of new maintenance technologies to expand the functionality and reliability of complete transformer substations are considered. Innovative functions of monitoring and intelligent control of automated systems for automatic input of switchgear reserve within the framework of the digital substation concept are proposed. Examples of new functions of the system of complete switchgears of the NKU "Cascade" based on HMI controllers using web technologies are given.

Keywords: digital substation, switchgear, automatic reserve input, maintenance, intellectualization.

УСТОЙЧИВОСТЬ ВЫСОКОВОЛЬТНОГО ЭЛЕКТРООБОРУДОВАНИЯ К НАРУШЕНИЯМ СИММЕТРИИ ПИТАЮЩИХ НАПРЯЖЕНИЙ

ООО «ТСН-электро», г. Нижний Новгород

Рассмотрены проблемы устойчивости автоматизированного электрооборудования по отношению к возмущениям в питающей сети как одному из наиболее вероятных видов возмущений. Проанализировано влияние данного вида возмущений на конкретном примере – высоковольтном регулируемом асинхронном электроприводе PowerFlex 7000. Предложены особенности реализации новых оригинальных технико-технологических решений функциональных возможностей электроприводных агрегатов с интеллектуальными системами управления и диагностики.

Ключевые слова: электрооборудование, электроприводной агрегат, системы электроснабжения, нарушения симметрии напряжений, устойчивость.

Введение

Возмущения в питающей энергосети оказывают существенное влияние на работу автоматизированных электроприводов (АЭП), вызывая необходимость применения сложных систем мониторинга [1-4]. Во многих случаях это влияние даже более существенно, чем влияние возмущений со стороны механизма и со стороны оператора – системы верхнего уровня иерархии [5-7]. На работу АЭП оказывают влияние отклонения модуля вектора напряжений, искажения напряжений, нарушение симметрии системы 3-фазных напряжений. Международным и отечественным стандартами в отношении качества питающих напряжений установлены нормально допустимые и предельные отклонения [8-11]. Предельные отклонения нормируются для вероятности 5 % при времени усреднения 24 час, что является на практике достаточно продолжительными отклонениями. Электропривод должен нормально работать с сохранением всех основных характеристик в условиях предельных отклонений. Однако в питающей сети возникают также аварийные отклонения.

Возможности АЭП в отношении устойчивости к аварийным отклонениям мало исследованы [12-14]. Поэтому уставки защит оборудования по отношению к питающей сети обычно недалеко выходят за рамки предельных отклонений, соответствующих стандартам, что зачастую приводит к неоправданному отключению оборудования и остановке технологического процесса. Нахождение допустимых границ уставок защит от возмущений питающей сети, при которых сохраняется работоспособность оборудования, – важная задача, которая поможет сократить потери, вызванные остановкой технологического процесса. В данной работе рассмотрен только один вид возмущения в питающей сети – нарушение симметрии питающего 3-фазного напряжения. Однако этот вид возмущений имеет наивысшую вероятность [15-17]. В этой связи актуальной задачей является разработка предложений в отношении показателей для оценки режимов при возмущении в сети с дальнейшим исследованием режимов на примере высоковольтного АЭП и выработке рекомендаций по выбору уставок защиты от нарушения симметрии.

Показатели для оценки режимов

Очевиден перечень показателей с убывающей приоритетностью:

1. Сохранение электропривода в работе.
2. Сохранение момента и скорости.
3. Сохранение в необходимых пределах пульсаций момента.
4. Сохранение в необходимых пределах искажений токов статора и токов, потребляемых от сети.

Рассмотрим векторный показатель искажения тока [18-19], наиболее полно учитывающий искажения. Для тока статора этот показатель определяется формулой:

$$I_{sD} = \sqrt{\frac{1}{T_M} \int_0^{T_M} |\mathbf{i}_s(t) - \mathbf{i}_{Fs}(t)|^2 dt}. \quad (1)$$

где T_M – время измерения, \mathbf{i}_s – изображающий вектор токов статора в системе координат, вращающейся синхронно, например, с полем двигателя, \mathbf{i}_{Fs} – фильтрованный вектор токов статора в этой же системе координат:

$$\mathbf{i}_{Fs}(t) = \frac{1}{1 + T_F p} \mathbf{i}_s(t). \quad (2)$$

Постоянная времени фильтра T_F также является одним из параметров измерения показателя искажения. Рассмотрим возмущения в сети, возникающие при однофазном (междуфазном) коротком замыкании на одном из ответвлений сети или при подключении к сети мощного двигателя с разрывом в цепи одной из фаз. Именно такой вид возмущений имеет наибольшую вероятность. Так в [11] указано, что 85 % провалов напряжения – именно такого вида. В таком режиме имеет место глубокий провал одного из линейных напряжений и некоторое снижение остальных линейных напряжений. Упрощенная расчётная схема и фазорная диаграмма представлены на рис. 1.

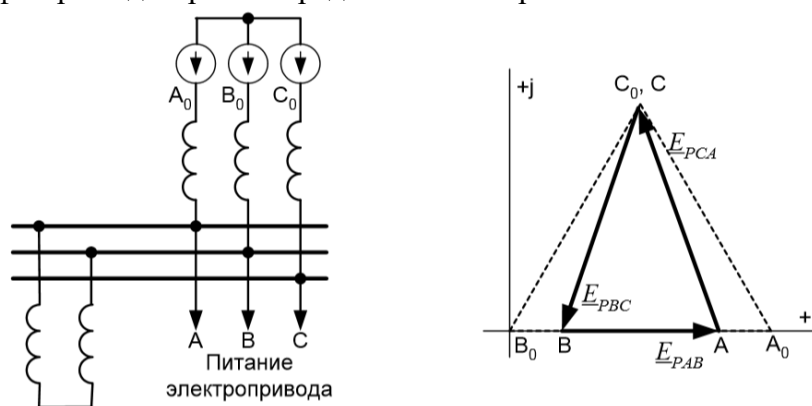


Рис. 1. Расчётная схема и фазорная диаграмма для нарушения симметрии

В данном режиме возникает напряжение обратной последовательности и несколько уменьшается напряжение прямой последовательности. Степень нарушения симметрии определяется относительным напряжением обратной последовательности U_{P2} / U_{PN} . Фазорная диаграмма рис. 1 не учитывает токов от регулируемого электропривода. Поэтому напряжения диаграммы могут рассматриваться как эквивалентные ЭДС трехфазного источника E_{PAB} , E_{PBC} , E_{PCA} , которым может быть представлена сеть по отношению к электроприводу.

Общая структура АЭП

В качестве примера исследован регулируемый асинхронный электропривод насоса, в состав которого входят: асинхронный двигатель и преобразовательно-регулирующее устройство (ПРУ) PowerFlex 7000 на основе инвертора тока с ШИМ, функциональная схема которого представлена на рис. 2. В состав ПРУ входят: мостовые преобразователи VUZL, VUZM на основе полностью управляемых полупроводниковых приборов, конденсаторные батареи фильтров CL, CM, сглаживающий реактор выпрямленного тока LD, трансформатор Т, устройство управления АС, а также элементы обратной связи по токам и напряжениям. Преобразователь VUZL преобразует входное напряжение в регулируемый выпрямленный ток; преобразователь VUZM преобразует выпрямленный ток в двуполярные пачки импульсов; высокочастотные составляющие этих пачек фильтруются конденсаторной батареей CM. Входные токи преобразователя VUZL также

представляют собой двуполярные пачки импульсов; высокочастотные составляющие фильтруются батареей CL.

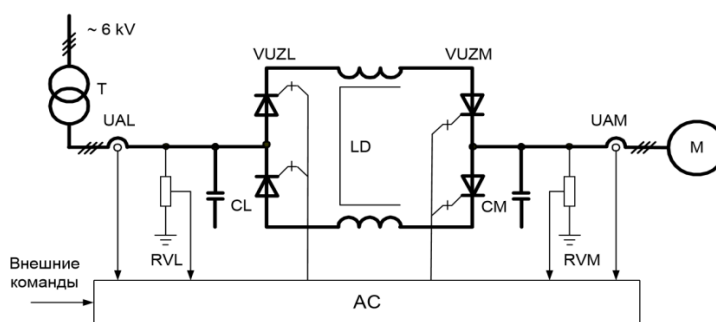


Рис. 2. Функциональная схема исследуемого электропривода

Воздействие нарушения симметрии питающих напряжений

Нарушение симметрии напряжений воздействует на данный АЭП по следующим направлениям:

1. В выпрямленном напряжении возникает дополнительная составляющая пульсаций с периодом, равным полупериоду питающих напряжений; соответственно увеличиваются пульсации выпрямленного тока. Это вызывает повышенное искажение токов статора и токов, потребляемых от сети, которое увеличивает пульсации момента.
2. При снижении линейных напряжений по расчётной схеме рис. 1 снижается постоянная составляющая выпрямленного напряжения. При достаточно высокой степени искажения сниженное выпрямленное напряжение не позволяет реализовать номинальный режим АЭП. САР снижает поток двигателя и уровень ограничения момента. Если АЭП работал с номинальной скоростью и полной нагрузкой, то при сильном искажении система снижает скорость электропривода.
3. Управление преобразователем VUZL синхронизировано с питающими напряжениями. При несовершенстве синхронизации и сильном нарушении симметрии питающих напряжений возможно нарушение синхронизации.

Параметры электропривода для нагрузки – насосе. Номинальные данные асинхронного двигателя: 1600 кВт, 1490 об/мин, 6 кВ, 190 А, 50 Гц. Электромеханическая постоянная времени электропривода $T_j = 0.8$ с. Относительные значения: постоянной составляющей момента нагрузки $M_{c0} = 0.1$, квадратичной составляющей момента при базовой скорости $M_{c2} = 0.7$.

Методика исследования на виртуальном стенде

Экспериментальное исследование выполнено на виртуальном испытательном стенде высоковольтных АЭП. В состав стенда входят шкаф управления серийного электропривода PowerFlex 7000 и компьютерная модель силовой части электропривода, работающая в реальном времени. Из шкафа управления в компьютерную модель поступают логические команды проводящего состояния для каждого плеча преобразователя частоты. Из модели в шкаф управления поступают аналоговые сигналы обратной связи по напряжениям и токам. В состав компьютерной модели входят модели механизма, двигателя, силовой части преобразовательного регулирующего устройства, питающей сети. Соответствие процессов стенда и реальных электроприводов было подтверждено многократно [20-21].

Все переменные и параметры, кроме времени и постоянных времени, представлены как относительные величины. В качестве базовых значений приняты:

- базовое напряжение – амплитуда номинального фазного напряжения двигателя $U_b = U_{phNm}$;
- базовый ток – амплитуда номинального фазного тока двигателя $I_b = I_{Nm}$;
- базовая угловая частота – номинальная угловая частота двигателя $\Omega_b = \Omega_N$;
- базовая угловая скорость – номинальная синхронная угловая скорость двигателя $\Omega_{MB} = \Omega_{MN}$.

Базовое значение тока трансформатора I_{Tb} связано с током I_b соотношением $I_{Tb} / I_b = U_N / U_{T2N}$, где U_{T2N} – номинальное вторичное напряжение трансформатора. Следует отметить, что при выбранных базовых значениях номинальный момент двигателя не совпадает с базовым моментом. Относительное значение номинального момента в рассматриваемом примере составляет $M_N = 0.816$.

Результаты исследования статических режимов

Исследованы режимы при скорости электропривода в пределах $v = 0.6-1$ при искажении питающих напряжений в пределах $E_{P2} / U_{PN} = 0-0,16$. Режимы при более сильном нарушении симметрии не рассматриваются: в таких случаях общее снижение линейных напряжений сети вызывает отключение электропривода защитой минимального напряжения с выдержкой времени 1-2 периода. На рис. 3, 4 в качестве примера представлены осциллограммы установившегося режима при скорости $v \approx 1$ и симметричном питании. Зарегистрированы переменные: i_D – выпрямленный ток, i_{FSA} – фазная компонента фильтрованного вектора тока статора, i_{SA} – фазный ток статора, M – электромагнитный момент, $|\Delta i_T|$ – модуль отклонения вектора первичного тока трансформатора от фильтрованного вектора, $|\Delta i_s|$ – модуль отклонения вектора тока статора от фильтрованного вектора. По данным, представленным на рис. 4, рассчитаны среднеквадратичные отклонения векторов тока статора и трансформатора I_{SD} , I_{TD} , а также среднеквадратичные отклонения выпрямленного тока и момента от их средних значений $\Delta i_{D,RMS}$, ΔM_{RMS} . Время измерения составляет $T_M = 5$ с, число точек на этом интервале $N = 50000$, постоянная времени для фильтров векторов тока $T_F = 1$ с. Такие же процессы, как представленный на рис. 4, зарегистрированы в указанных ранее диапазоне скорости и диапазоне нарушения симметрии питающих напряжений. В табл. 1 представлены результаты измерений для номинальной скорости $v \approx 1$.

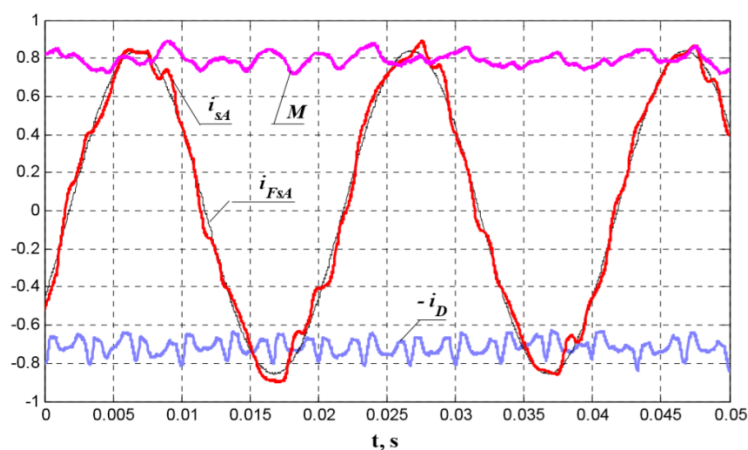


Рис. 3. Установившийся режим при симметричном питании

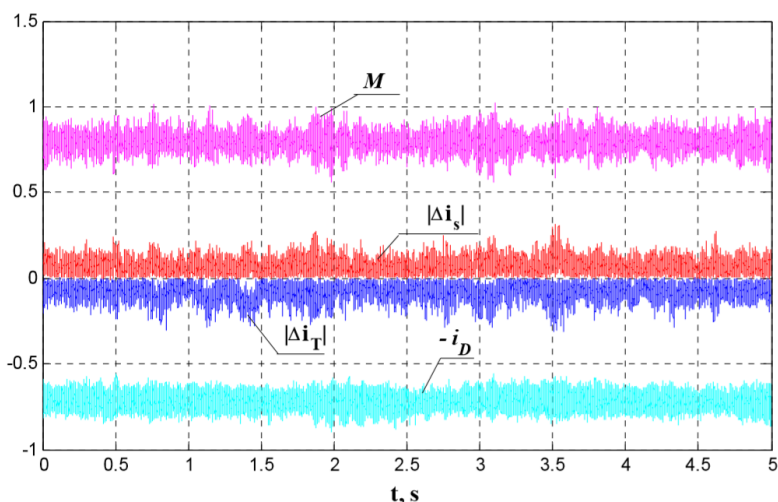


Рис. 4. Осциллограмма для расчёта среднеквадратичных отклонений

Таблица 1

Режимы электропривода при нарушении симметрии питающих напряжений

Показатели режима	Напряжение обратной последовательности E_{P2} / U_{PN}			
	0	0.04	0.10	0.16
Среднее значение момента M_{av}	0.80	0.80	0.79	0.71
Среднеквадратичное отклонение выпрямленного тока $\Delta i_{D,RMS}$	0.046	0.047	0.063	0.084
Среднеквадратичное отклонение вектора тока трансформатора I_{TD}	0.099	0.106	0.179	0.230
Среднеквадратичное отклонение вектора тока статора I_{sD}	0.091	0.094	0.126	0.155
Среднеквадратичное отклонение момента ΔM_{RMS}	0.059	0.060	0.079	0.106

Измерения показывают, что режим с нарушением симметрии $E_{P2} / U_{PN} = 0.04$, предельно допустимый по отечественному стандарту, практически не отличается от режима с симметричным питанием; АЭП может продолжительно работать в таком режиме. При увеличении напряжения обратной последовательности до значения $E_{P2} / U_{PN} = 0.1$ среднеквадратичное отклонение выпрямленного тока возрастает примерно в 1.35 раза. Примерно в таком же отношении возрастают среднеквадратичные отклонения вектора тока статора и момента. Среднеквадратичное отклонение вектора тока трансформатора возрастает в 2.3 раза. Электропривод сохраняет момент и скорость. Он в состоянии продолжать работу в течение достаточного времени, пока не восстановится нормальный режим системы электроснабжения. При значительно более сильном нарушении, при $E_{P2} / U_{PN} = 0.16$, существенно нарастают искажения токов и момента. Однако под действием системы регулирования снижается скорость АЭП, что отражается в снижении среднего момента. Минимальное линейное напряжение сети в этом режиме составляет всего 68 % номинала. Этого недостаточно для поддержания номинальной скорости при номинальном магнитном потоке двигателя. Режим АЭП не является аварийным, однако он допустим лишь на несколько секунд. Работа АЭП при пониженной скорости имеют такие же особенности, но снижение скорости при повышенном искажении не происходит.

Рекомендации по уставкам защиты от нарушения симметрии.

По результатам исследований возможно задать уровень срабатывания защиты от нарушения симметрии, соответствующий напряжению обратной последовательности $U_2 = 0.1 U_{PN}$. Выдержка времени защиты может быть установлена в пределах $T_{UV} = 1-5$ с.

Выводы

Устойчивость электропривода по отношению к возмущениям в питающей сети является одной из важнейших характеристик. Однако данные по противодействию аварийным возмущениям не приводятся для высоковольтных АЭП большой мощности. АЭП может работать продолжительно при напряжении обратной последовательности в 2.5 раза превышающем предельно допустимое значение по ГОСТ. Результаты исследования позволяют изменить уставки соответствующей защиты электропривода и сохранять АЭП в работе при более сильных нарушениях, например, 3.15 МВт шахтных вентиляторов рудника.

Библиографический список

1. Васенин А.Б., Степанов С.Е., Крюков О.В. Методология и средства оперативного мониторинга электродвигателей на КС // Контроль. Диагностика. 2019. № 11. С. 52-58.
2. Крюков О.В. Мониторинг условий эксплуатации электродвигателей газоперекачивающих агрегатов // Контроль. Диагностика. 2016. № 12. С. 50-58.
3. Мещеряков В.Н., Ласточкин Д.В., Крюков О.В. Приложения теории нечетких множеств для обработки данных и задач прогнозирования в системах АЭП // В сборнике: Современные сложные системы управления. Материалы XII МНПК. 2017. С. 153-158.
4. Крюков О.В., Серебряков А.В. Активно-адаптивные алгоритмы управления и мониторинга автономными энергетическими комплексами // В сборнике: Пром-Инжиниринг. Труды II МНТК. ЮУрГУ. 2016. С. 286-290.
5. Крюков О.В., Степанов С.Е. Модернизация систем управления ЭГПА в условиях действующих компрессорных станций // В сборнике: Проблемы автоматизации и управления в технических системах. МНТК под редакцией М.А. Щербакова. 2013. С. 29-32.
6. Крюков О.В. Комплексный анализ условий эксплуатации электродвигателей ГПА // Компрессорная техника и пневматика. 2013. № 4. С. 14-19.
7. Кононенко А.Б., Косоротов А.А., Крюков О.В. Расширение функциональных возможностей автоматизации и мониторинга распределительных устройств КТП «Каскад» // Автоматизация и ИТ в энергетике. 2020. № 12. С. 26-31.
8. Реализованные проекты ООО «ТСН-электро». Режим доступа: <https://www.tcn-nn.ru/>
9. International Standard IEC 61000-2-12. Electromagnetic compatibility (EMC) – Part 2-12: Environment – Compatibility levels for low-frequency conducted disturbances and signaling in public medium-voltage power supply systems. – 2003.04.
10. ГОСТ 13109-1997. Электрическая энергия. Совместимость технических средств электромагнитная. Нормы качества электрической энергии в системах электроснабжения общего назначения (переиздание, январь 2002).
11. Power Quality Newsletter – May 2007. Power Standards Lab., California, USA. – <http://www.powerstandards.com>.
12. Крюков О.В. Алгоритмы быстрого преобразования Уолша в микропроцессорных системах управления электроприводом // Известия ВУЗов. Электромеханика. 2005. № 4. С. 39-44.

13. Kryukov O.V., Gulyaev I.V., Teplukhov D.Y. Method for stabilizing the operation of synchronous machines using a virtual load sensor // Russian Electrical Engineering. 2019. T. 90. № 7. С. 473-478.
14. Kryukov O.V., Blagodarov D.A., Dulnev N.N., al. Intelligent control of electric machine drive systems // В сборнике: 2018 10th International Conference on Electrical Power Drive Systems, ICEPDS 2018 - Conference Proceedings 10. 2018. С. 8571670.
15. Крюков О.В. Опыт проектирования АСУ ТП нефтеперекачивающих станций магистральных нефтепроводов // Приборы и системы. Управление, контроль, диагностика. 2017. № 1. С. 2-7.
16. Belousov A.S., Meshcheryakov V.N., Valtchev S., Kryukov O.V. Development of a control algorithm for three-phase inverter in two-phase electric drives reducing the number of commutations // В сборнике: Proceedings - 2019 1st International Conference on Control Systems, Mathematical Modelling, Automation and Energy Efficiency, SUMMA 2019. 2019. С. 444-449.
17. Крюков О.В., Степанов С.Е., Серебряков А.В. Современный подход к организации ремонта по данным прогноза технического состояния и ресурса электрооборудования // Газовая промышленность. 2017. № 8 (756). С. 84-89.
18. Крюков О.В., Мещеряков В.Н., Гуляев И.В. Электроприводы на основе машины двойного питания и асинхронного вентильного каскада с преобразователями в цепях статора и ротора // Саранск, 2020.
19. Vasenin A.B., Kryukov O.V., Serebryakov A.V. Adaptive control algorithm sofautonomous generator complexes // В книге: Электромеханика, электротехнологии, электротехнические материалы и компоненты. Труды МКЭЭЭ-2016. 2016. С. 133-135.
20. Васенин А.Б., Крюков О.В. Энергоэффективные и экологичные установки воздушного охлаждения // В сборнике: ВЕЛИКИЕ РЕКИ' 2017 труды научного конгресса 19-го Международного научно-промышленного форума. НГАСУ. 2017. С. 93-96.
21. Васенин А.Б., Крюков О.В. Проектирование электромеханической части и систем управления энергетических установок газотранспортных потребителей // Известия Тульского гос. ун-та. Технические науки. 2011. № 5-1. С. 47-51.

V.A. Ippolitov, D.V. Proskuryakov, O.V. Kryukov

STABILITY OF HIGH-VOLTAGE ELECTRICAL EQUIPMENT TO SUPPLY VOLTAGE SYMMETRY VIOLATIONS

TSN-electro LLC, Nizhny Novgorod

The problems of stability of automated electrical equipment in relation to disturbances in the supply network as one of the most likely types of disturbances are considered. The influence of this type of perturbation is analyzed on a specific example – a high-voltage regulated asynchronous electric drive PowerFlex 7000. The features of the implementation of new original technical and technological solutions of the functional capabilities of electric drive units with intelligent control and diagnostic systems are proposed.

Keywords: electrical equipment, electric drive unit, power supply systems, stress symmetry violations, stability.

ВОЗМОЖНОСТИ ПОВЫШЕНИЯ ИСКРОБЕЗОПАСНОСТИ ЭЛЕКТРООБОРУДОВАНИЯ ОБЪЕКТОВ ТЭК

ООО «ТСН-электро», г. Нижний Новгород

Рассмотрены пути повышения искробезопасности ответственного электрооборудования, работающего в нефтегазовой, химической и перерабатывающей отраслях промышленности. Проведен анализ отечественной нормативной базы и международных стандартов в области защиты силового оборудования и измерительных цепей в части защиты от внешних и внутренних перенапряжений. Показана необходимость и даны рекомендации по использованию УЗИП в топливно-энергетическом и нефтегазовом комплексе.

Ключевые слова: искробезопасность, искрозащита, устройство защиты от импульсных перенапряжений, электроприводной газоперекачивающий агрегат.

Введение

Для современных энергетических предприятий важнейшей целью является не только производство продукции высокого качества, но и обеспечение надежной и безаварийной работы всех технологических цепочек на высоком технико-экономическом уровне, отвечающем современным требованиям промышленной безопасности [1-3].

Предприятия нефтегазового сектора, энергетической и горнодобывающей отраслей промышленности характеризуются наличием объектов повышенной опасности (ОПО), взрывоопасных зон, на которых требуется осуществлять непрерывный мониторинг [4-6] с прогнозированием технического состояния в среднесрочной перспективе [7-8]. Надежность и оперативность таких систем играют важнейшую роль в непрерывном производственном процессе, например, компрессорных станций (КС) магистральных газопроводов (МГ) [9-12]. Электрическая инфраструктура всех ОПО очень чувствительна к импульсным помехам, которые способны ограничить работоспособность промышленных установок и в целом производственных объектов по сетям питания и информационным каналам [13-14].

Устройства молниезащиты и защиты от импульсных перенапряжений (УЗИП) призваны обеспечить защиту таких ОПО КС, как электроприводной газоперекачивающий агрегат (ЭГПА) со сложной и дорогостоящей техникой, распределительные устройства (РУ) комплектных трансформаторных подстанций (КТП) различных классов напряжения (рис. 1), серверное и цифровое оборудование систем электроснабжения [15-16]. Для защиты оборудования и всей системы производства в целом и, в частности, предотвращения травм персонала, необходимо, чтобы концепция молниезащиты и защиты от импульсных перенапряжений предприятия отвечала всем действующим стандартам [17-21], что одинаково верно, как для взрывоопасных производственных зон ОПО, так и для взрывобезопасных зон.

В настоящее время производимые ООО «ТСН-электро» НКУ предназначены для распределения электроэнергии трехфазного переменного тока напряжением до 690 В частотой 50 Гц в сетях с изолированной и глухозаземленной нейтралью, для защиты линий при перегрузках и коротких замыканиях, а также для управления (регулирования, автоматике), измерения, сигнализации и защиты оборудования.

Новые системы НКУ типа «Каскад» могут применяться в качестве РУ во всех сферах энергопотребления, где требуется обеспечить ввод и распределение электрической энергии, в частности: в системе собственных нужд всех типов электростанций; для комплектования подстанций электрических сетей; для комплектования подстанций перекачивающих станций нефтегазопроводов; в системах электроснабжения и автоматики промышленных предприятий и коммунальной сферы.



Рис. 1. Современное электротехническое оборудование НКУ «Каскад» для КТП

Принципы построения и нормативные документы по УЗИП

Для решения указанной задачи во взрывоопасных зонах каналы связи оборудуются искробезопасными цепями. Для сопряжения электрооборудования, расположенного во взрывоопасной зоне и вне ее, используются ограничительные устройства - *барьеры искробезопасности*, обеспечивающие искрозащиту электрических цепей датчиков, находящихся во взрывоопасной зоне. Устройства применяются в системах автоматического регулирования, сигнализации и аварийной защиты на взрыво-, пожароопасных участках, где могут присутствовать взрывоопасные смеси газов, паров, а также легковоспламеняющиеся и взрывчатые вещества.

Безопасность в нефтегазовой, химической и перерабатывающей промышленности должна отвечать самым высоким требованиям. В частности, директива АТЕХ, и соответствующие отраслевые и международные стандарты включают многочисленные инструкции, которые должны быть скоординированы для практического применения. Для новых предприятий это реализуется еще на этапе планирования производства. Директива АТЕХ рассматривает оценку опасностей взрыва и возможных источников зажигания, в их числе помимо прочих упоминаются удары молний, коммутационные всплески и электростатические разряды. Эти же факторы указываются в Правилах устройства электроустановок (ПУЭ 7). Устройства для защиты от импульсных перенапряжений, вызванные перечисленными факторами, регламентируются действующим стандартом МЭК 61643-1. В России данный стандарт издан в виде ГОСТ Р 51992-2002. Методы защиты от физических повреждений зданий и сооружений, вызванных грозовыми разрядами, регламентируются международным стандартом МЭК 62305. При этом само электрооборудование, установленное во взрывоопасной зоне должно соответствовать следующим стандартам:

- МЭК 60079-0 — Электрооборудование для взрывоопасных газовых сред;
- МЭК 61241-10 — Электрооборудование для сред с горючей пылью.

Данные стандарты также гармонизированы со стандартами ГОСТ Р.

Оценка риска выполняется для того, чтобы определить необходимость в системе молниезащиты и выбрать защитные меры приемлемые и с технической, и экономической точек зрения. Система молниезащиты включает внешнюю систему молниезащиты, токоотводящие шины, систему заземления, а также внутреннюю защиту от импульсных перенапряжений. В стандарте МЭК 62305 обсуждается обширный перечень факторов с целью калькуляции возможных убытков, а также чтобы найти пути их избежать и минимизировать риски. Концепция молниезащиты в целом и защита от импульсных перенапряжений в частности должна использоваться в обязательном порядке во взрывоопасном производстве, т.к. возможные последствия этих факторов становятся значительно более критичными как для оборудования, так и для людей.

В приложении к части 3 стандарта МЭК 62305 содержится важная информация о практической реализации и деталях системы заземления и системы выравнивания потенциалов. Выравнивание потенциалов системы молниезащиты представлена, как часть стандарта под пунктом 6.2. Это достигается соединением следующих компонентов, используя проводники выравнивания потенциалов, естественные соединения, устройства защиты от импульсных перенапряжений (УЗИП):

- металлические части конструкций;
- металлические установочные компоненты;
- внешние проводящие линии, имеющие соединения с аппаратной частью системы;
- электрические и электронные системы с аппаратной частью, которая должна быть защищена.

Цель молниезащиты и защиты от импульсных перенапряжений — повышение безопасности производства и его готовности. Принципы безопасного оборудования системы для перерабатывающей промышленности с привязкой к ЭМС, так же, как и опасности, вызванные ударами молний, собраны в стандарте МЭК 61511-1.

Зонная концепция молниезащиты формирует основу системы молниезащиты (рис. 2).

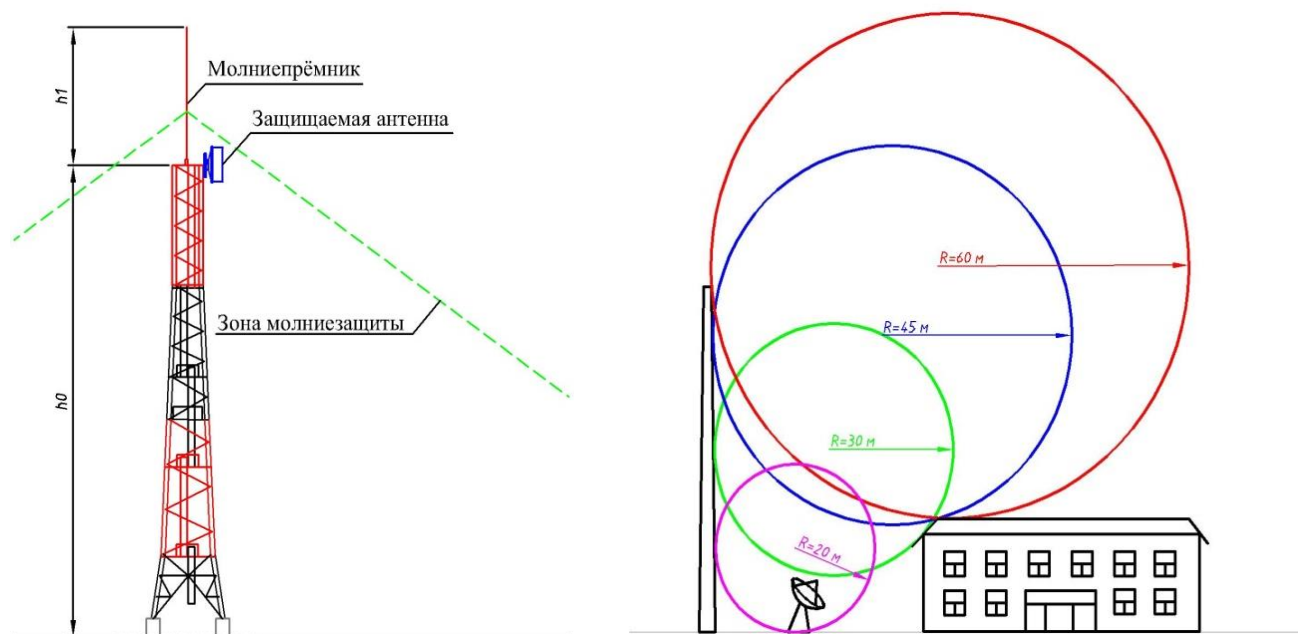


Рис. 2. Примеры защиты АФУ по СТО Газпром с применением метода фиктивной сферы

Ее описание приведено в международном стандарте МЭК, перенесено в российскую «Инструкцию по устройству молниезащиты зданий, сооружений и промышленных коммуникаций СО 153-34.21.122-2003», а также повторяется во многих локальных национальных отраслевых стандартах, в том числе ПАО «Газпром».

В соответствии с этой концепцией оборудование классифицируется в соответствии с зонами молниезащиты (LPZ — lightning protection zones) от 0А до 3.

Целью является последовательное уменьшение электромагнитных возмущений от зоны 0А до зоны 2 или 3, вызванных близким или прямым ударом молнии, до допустимого для каждого конкретного случая уровня. Внешние элементы молниезащиты и механические части, окружающие полевые приборы, должны предотвращать оборудование от разрушений, вызванных прямым попаданием разряда молнии. Защита от импульсного перенапряжения должна защищать электрическое и электронное оборудование от электромагнитных воздействий, вызванных разрядом молнии.

Особенности применения УЗИП в газовой промышленности

В нефтегазовой, химической и перерабатывающей промышленности, где производственные площадки разбросаны на очень больших территориях с длинными линиями сетей передачи данных и питающими линиями, в электрической инфраструктуре появляются значительные разности потенциалов во время грозы при ударах молнии в объекты или близлежащие территории. Также импульсное перенапряжение появляется при коммутации индуктивной нагрузки, такой, как электродвигатели. В результате повреждения оборудования может произойти короткое замыкание в сети или пробой на землю и, как следствие, отключаются целые секции электрической системы.

Также эксплуатационные предприятия должны принимать в расчет последствия аварийного отключения электрооборудования, когда оборудование отключается под нагрузкой, коммутационные операции этого типа могут вызвать возникновение импульсных перенапряжений в несколько тысяч вольт. Даже если в результате не последует немедленного выхода из строя системы, электрооборудование может быть повреждено, что послужит выходу системы из строя через некоторое время. Впоследствии достаточно трудно установить, что причиной этого было воздействие импульсного напряжения из-за удара молнии или аварийного отключения оборудования неделю назад.

УЗИП позволяют существенно повысить степень готовности производственных систем, а зонная концепция защиты облегчает выбор решения. Соответствующие меры защиты от импульсных перенапряжений и последовательное уравнивание потенциалов должны быть приняты на переходах из одной защитной зоны в другую. Для цепей питания принято устанавливать УЗИП классов 1, 2 и 3 на границах соответствующих зон.

Устройства защиты от импульсных перенапряжений для измерительных и информационных цепей — токовые петли, цифровые сигналы, магистральные системы — имеют конструкцию, включающую как минимум два каскада защиты, а, следовательно, обеспечивают защиту устройств при передаче сигнала через разные защитные зоны. Они обладают достаточно высокой способностью отводить разрядные токи и при этом имеют низкий уровень защиты (напряжение, при котором происходит срабатывание устройства) и маленькое время срабатывания.

В измерительных линиях, следует принять во внимание, что в защите нуждаются с одной стороны полевые устройства, а с другой — устройства контроля и управления. УЗИП для измерительных цепей, как известно, работают по принципу предохранительного короткого замыкания. В случае перегрузки они замыкают сигнальную цепь на линию защитного заземления, чем предотвращают повреждение соответствующего интерфейса.

С точки зрения экономической эффективности производства, одним из важнейших вопросов является стоимость эксплуатации, ремонта и времени простоя системы. Опыт показывает, что предприятия, применяющие в своих системах УЗИП, значительно повышают степень надежности своего оборудования. Поэтому в ОПО помимо повышения уровня экономической эффективности, огромное значение имеет уровень безопасности всего производства.

Компенсация опасных стохастических возмущений на электрооборудование

Совокупность случайных факторов, действующих на рабочие органы агрегатов, определяет не только случайный характер нагрузки, потребляемой мощности электрооборудование агрегата, но и его КПД, а также коэффициент мощности. Так как случайный процесс нагружения не может быть достоверно описан традиционной детерминированной диаграммой, для адекватного расчета основных параметров технологического электрооборудования в этом случае необходимо использовать статистические методы теории планирования эксперимента. Исследования с привлечением аппарата теории вероятности показали, что процесс нагружения приводов обычно распределен по закону Райса.

При отсутствии непосредственного воздействия совокупности стохастических возмущений на электрические и механические органы агрегатов (системы 2-го рода) наиболее целесообразным является организация инвариантной системы управления, адекватно учитывающей все частные влияния возмущений на объект. Для этого все влияющие возмущения после формализации вводятся в закон управления, обеспечивая адекватность задания момента и скорости текущему состоянию объекта. При этом реализуются замкнутые САУ по возмущению, отклонению и комбинированные.

Помимо основных ЭГПА характерными примерами агрегатов КС со случайными возмущениями 2-го рода являются локальные приводы, входящие в АСУ КС: аппараты воздушного охлаждения (АВО) газа после его компримирования в ЭГПА, АВО масла всех основных систем и агрегатов КС, а также насосы и вентиляторы вспомогательных систем, параметры которых подвержены климатическим воздействиям. Для рассматриваемой группы механизмов и технологических процессов возможны два метода идентификации и расчета, используемые при проектировании АСУ ТП.

1. Идентификация с последующим автоматизированным расчетом алгоритмов управления. В этом случае идентификация модели объекта управления осуществляется один раз, после чего рассчитывается алгоритм управления с постоянными параметрами в режиме *on-line* или *off-line*. Преимуществами метода является возможность расчета и оценки любого алгоритма с дальнейшим моделированием и созданием общей модели. Поэтому данный метод следует применять при выборе структуры системы управления и расчета регуляторов с фиксированными параметрами или адаптивных регуляторов с прямой связью.

2. Разработка самооптимизирующихся (самонастраивающихся) адаптивных алгоритмов управления агрегатами. В этом случае идентификация модели объекта производится периодически и после очередного получения оценок модели объекта в режиме *on-line* определяются параметры алгоритма управления с переменными коэффициентами. Достоинствами метода является возможность непрерывного слежения алгоритма за медленно меняющимися параметрами объекта и меньшее время обработки в ПК.

Анализ требований показывает, что все они могут быть реализованы на основе современного серийного электрооборудования. Однако реализация требования по стохастическим возмущениям предполагает применение для их обработки и формализации статистических методов, основанных на приложениях центральной теоремы теории вероятности, входящих в общую теорию планирования эксперимента [3,9,16].

Выводы

Проведен анализ отечественной нормативной базы и международных стандартов в области защиты силового оборудования и измерительных цепей в части защиты от внешних и внутренних перенапряжений. Доказана необходимость и даны рекомендации по использованию УЗИП в топливно-энергетическом и нефтегазовом комплексе.

Библиографический список

1. Крюков О.В., Степанов С.Е. Повышение искробезопасности электрооборудования в газовой промышленности // Охрана труда и техника безопасности на промышленных предприятиях. 2014. № 8. С. 18-21.
2. Крюков О.В. Комплексный анализ условий эксплуатации электродвигателей ГПА // Компрессорная техника и пневматика. 2013. № 4. С. 14-19.
3. Васенин А.Б., Крюков О.В. Проектирование электромеханической части и систем управления энергетических установок газотранспортных потребителей // Известия Тульского гос. ун-та. Технические науки. 2011. № 5-1. С. 47-51.
4. Крюков О.В. Мониторинг условий эксплуатации электродвигателей газоперекачивающих агрегатов // Контроль. Диагностика. 2016. № 12. С. 50-58.
5. Крюков О.В., Серебряков А.В. Активно-адаптивные алгоритмы управления и мониторинга автономными энергетическими комплексами // В сборнике: Пром-Инжиниринг. Труды II МНТК. ЮУрГУ. 2016. С. 286-290.
6. Васенин А.Б., Степанов С.Е., Крюков О.В. Методология и средства оперативного мониторинга электродвигателей на КС // Контроль. Диагностика. 2019. № 11. С. 52-58.
7. Мещеряков В.Н., Ласточкин Д.В., Крюков О.В. Приложения теории нечетких множеств для обработки данных и задач прогнозирования в системах АЭП // В сборнике: Современные сложные системы управления. Материалы XII МНПК. 2017. С. 153-158.
8. Крюков О.В., Степанов С.Е., Серебряков А.В. Современный подход к организации ремонта по данным прогноза технического состояния и ресурса электрооборудования // Газовая промышленность. 2017. № 8 (756). С. 84-89.
9. Васенин А.Б., Крюков О.В. Энергоэффективные и экологичные установки воздушного охлаждения // В сборнике: ВЕЛИКИЕ РЕКИ' 2017 труды научного конгресса 19-го МНПФ. НГАСУ. 2017. С. 93-96.
10. Крюков О.В., Степанов С.Е. Модернизация систем управления ЭГПА в условиях действующих компрессорных станций // В сборнике: Проблемы автоматизации и управления в технических системах. МНТК под редакцией М.А. Щербакова. 2013. С. 29-32.
11. Kryukov O.V., Serebryakov A.V. Modern systems of outdoor illumination for compressor stations // Light & Engineering. 2016. Т. 24. № 2. С. 128-131.
12. Захаров П.А., Крюков О.В. Перспективы развития метрологических расходоизмерительных центров ПАО "Газпром" // В сборнике: Информационные системы и технологии ИСТ-2018. Мат. XXIV МНТК к 100-летию Нижегородской радиолоборатории. 2018. С. 693-699.
13. Vasenin A.B., Kryukov O.V., Serebryakov A.V. Adaptive control algorithms of autonomous generator complexes // В книге: Электромеханика, электротехнологии, электротехнические материалы и компоненты. Труды МКЭЭЭ-2016. 2016. С. 133-135.
14. Крюков О.В. Алгоритмы быстрого преобразования Уолша в микропроцессорных системах управления электроприводом // Известия ВУЗов. Электромеханика. 2005. № 4. С. 39-44.
15. Васенин А.Б., Крюков О.В. Вопросы электропитания вдоль трассовых объектов ЕСГ России // Научно-технический сборник Вести газовой науки. 2020. № 2 (44). С. 181-192.
16. Васенин А.Б., Хлынин А.С., Крюков О.В. Автоматизированные системы собственных нужд КС // В сборнике: Информационные системы и технологии ИСТ-2020 XXVI МНТК. НГТУ им. Р.Е. Алексеева. 2020. С. 293-299.
17. ГОСТ 9.005-72. Допустимые и недопустимые контакты металлов.

18. ГОСТ Р 50571.21-2000. (МЭК 60364-5-548-96). Заземляющие устройства и системы уравнивания электрических потенциалов в электроустановках, содержащих оборудование обработки информации.
19. ГОСТ 10434-82. Соединения контактные электрические. Классификация. Общие технические требования.
20. ГОСТ 9.303-84. Единая система защиты от коррозии и старения. Покрытия металлические и неметаллические неорганические. Общие требования к выбору и обозначения.
21. Терентьев Д.Е. Минимизация и стабилизация переходного сопротивления разъемных контактов в системах уравнивания потенциалов и заземляющих устройствах / Состояние и перспективы развития энергетики и связи - www.commmeng.ru, www.spres.info.

B. Bugrezov, A. E. Putera, O. V. Kryukov

OPPORTUNITIES FOR IMPROVING THE INTRINSIC SAFETY OF ELECTRICAL EQUIPMENT OF FUEL AND ENERGY FACILITIES

TSN-electro LLC, Nizhny Novgorod

The ways of improving the intrinsic safety of responsible electrical equipment operating in the oil and gas, chemical and processing industries are considered. The analysis of the domestic regulatory framework and international standards in the field of protection of power equipment and measuring circuits in terms of protection against external and internal overvoltages is carried out. The necessity is shown and recommendations are given for the use of the UZIP in the fuel and energy and oil and gas complex.

Keywords: intrinsic safety, spark protection, surge protection device, electric gas pumping unit.

ДИНАМИКА ЧАСТОТНОГО ПУСКА АВТОМАТИЗИРОВАННОГО ЭЛЕКТРОПРИВОДА

ООО «Газпром проектирование», г. Н. Новгород,

МГУ им. Н. П. Огарёва, г. Саранск,

ООО «ТСН-электро», г. Нижний Новгород

Рассмотрены высокоэффективные методы частотного запуска двигателей основных технологических агрегатов ТЭК. Анализ схмотехнических вариантов частотного пуска показал возможности значительного улучшения всех основных недостатков прямого пуска. Представлены преимущества автоматической синхронизации с питающей сетью и автоматическое безударное переключение статора двигателя на сеть. Наилучшие практики гладкого переключения на сеть обеспечивается только в варианте с алгоритмом работы без паузы при питании цепи статора с минимумом перерегулирования по току статора и колебаний электромагнитного момента двигателя.

Ключевые слова: частотный пуск, электродвигатель, пусковое регулирующее устройство, пауза, переключения, моделирование.

Введение

Частотный пуск используется для двигателей переменного тока большой мощности, в тех случаях, когда не требуется регулирование скорости приводимого механизма [1-4]. Частотный пуск устраняет неблагоприятные факторы, имеющие место при прямом пуске, в отношении воздействия на двигатель, механизм и питающую сеть [5-8].

Обобщённая функциональная схема системы частотного запуска электроприводного агрегата приведена на рисунке 1 [9-12].

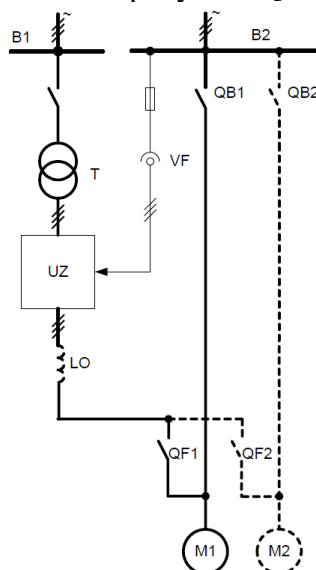


Рис. 1. Функциональная схема системы частотного запуска электроприводного агрегата

На схеме показаны: электродвигатель $M1$, преобразовательное регулирующее устройство (ПРУ) UZ , трансформатор T , элемент обратной связи по напряжению VF , контакторы (выключатели) $QF1$, $QB1$, секции распределительного устройства (РУ) $B1$, $B2$. Последнее может быть интегрировано в одну и ту же секцию. При необходимости в некоторых случаях используется дополнительный реактор LO на выходе ПРУ. Одно преобразовательное регулирующее устройство часто используют для частотного пуска нескольких двигателей [13-16]. Часть схемы, относящаяся к двигателю $M2$, показана на

рис. 1 пунктирными линиями. Общий принцип частотного пуска состоит в том, что двигатель вначале подключается (контактором $QF1$) к выходу ПРУ, ускоряется в составе регулируемого электропривода до определённой скорости, близкой к синхронной, и синхронизируется с сетью, на которую он должен переключиться. После этого статор двигателя отключается от ПРУ и подключается к сети (контактором QB). Процедура синхронизации на сеть называется синхронным трансфером, основная задача которого – обеспечить плавное переключение, без заметных толчков тока статора и момента. Используются виды синхронного трансфера, отличающиеся последовательностью работы контакторов $QF1$, $QB1$ и методом синхронизации с сетью.

Переключение при работающем ПРУ

Когда частота напряжения статора в процессе ускорения приближается к частоте сети, замыкаются дополнительные контуры регулирования [17-20]: контур регулирования углового положения вектора напряжения статора и контур модуля вектора напряжения статора. Эти контуры стремятся свести к нулю рассогласование векторов напряжения статора и сети. После того, как погрешности по угловому положению и по модулю входят в заданные пределы, статор подключается к сети контактором $QB1$. Затем уже запирается ПРУ и отключается контактор $QF1$. Метод обеспечивает практически идеальное переключение.

Однако недостатком метода является подключение выхода работающего ПРУ к сети, и он в течение определённого времени остаётся подключённым к сети. Это создаёт угрозу сверхтоков в силовых контурах ПРУ по двум возможным причинам: ошибка в начальной синхронизации и повреждение внутри ПРУ на интервале его совместной работы с сетью.

При большой мощности короткого замыкания на шинах, к которым переключается двигатель, мгновенные значения сверхтоков могут превышать значения, допустимые по условиям даже не повреждения, а разрушения силовых полупроводниковых элементов в составе ПРУ. Для каждого такого элемента в каталоге производителя указывается максимальное мгновенное значение тока по условиям разрушения (*SurgeCurrent*). При потенциальной возможности таких токов на выходе ПРУ включается токоограничивающий реактор LO (рис. 1).

Переключение с ПРУ на сеть с паузой

По этому методу вначале запирается ПРУ. Только после этого статор подключается к сети. Имеет место пауза продолжительностью T_p , в течение которой статор двигателя отключён от каких-либо источников. К концу паузы скорость вращения вектора напряжения статора становится равной скорости ротора. Модуль вектора напряжения статора к этому моменту определяется потоком ротора АД и основным магнитным потоком СД. За время паузы скорость ротора снизится на величину $\Delta\omega_r$, и произойдёт отставание вектора потока по углу на величину $\Delta\gamma_r$.

Последовательность работы контакторов $QF1$, $QB1$ может быть различной в различных вариантах данного метода. Возможен, например, вариант, в котором контактор $QB1$ включается раньше, чем отключается контактор $QF1$. В этом случае во время паузы статор двигателя подключён к выходу отключенного ПРУ. Если в составе ПРУ имеется выходной фильтр, например, в преобразователе частоты (ПЧ) на основе инвертора тока (ИТ) с ШИМ и выходной конденсаторной батареей, то статор во время паузы подключён к конденсаторной батарее фильтра.

В любом случае, после запираания ПРУ происходит очень быстрое и весьма значительное изменение вектора тока статора. Также большие изменения претерпевает и вектор потокосцепления статора. Если даже в момент подключения к сети вектор эквивалентного потокосцепления сети совпадает с тем значением, которое он должен иметь в новом установившемся режиме, то установившийся режим не может возникнуть мгновенно. Поэтому неизбежен электромагнитный и электромеханический переходный процесс.

В результате при наличии паузы принципиально невозможно идеальное переключение. Неизбежны изменения в токе статора, в моменте и в скорости двигателя. Основная задача состоит только в том, чтобы исключить сверхтоки статора и свести к минимуму колебания момента.

Переключение без начального опережения по скорости представляет собой достаточно простой метод. В процессе ускорения двигателя при подходе к номинальной скорости замыкается контур регулирования фазы вектора потокосцепления статора. Вектор потокосцепления статора сравнивается по фазе с вектором эквивалентного потокосцепления сети. Задаётся угол опережения вектора потокосцепления статора. Регулятор фазы вводит дополнительное задание на вход регулятора скорости.

Желательно также обеспечить регулирование модуля вектора потокосцепления статора воздействием на регулятор магнитного потока. После того, как регулятор фазы вводит систему в установившийся режим, начинаются операции по переключению. Подается команда на включение контактора $QB1$. Поскольку время включения этого контактора довольно значительно, с задержкой запирается ПРУ. Задержка программируется таким образом, чтобы обеспечивалась пауза с заданным малым временем. При использовании стабильного выключателя удаётся работать с паузой порядка 36-40 ms .

На рис. 2 представлена осциллограмма процессов трансфера для электропривода нагнетателя газа мощностью 3 MW , 1490 об/мин. Схема ПРУ – на основе ИТ с ШИМ. Электромеханическая постоянная времени электропривода $T_j = 13.6 s$. Представлены процессы нагруженного привода. Время паузы $T_P = 40 ms$.

Зарегистрированные переменные: i_D – выпрямленный ток ПРУ, i_s – ток статора (модуль вектора), u_s – напряжение статора (модуль вектора). Во время паузы ток статора не падает до нуля – статор остаётся подключённым к конденсаторной батарее фильтра ПРУ.

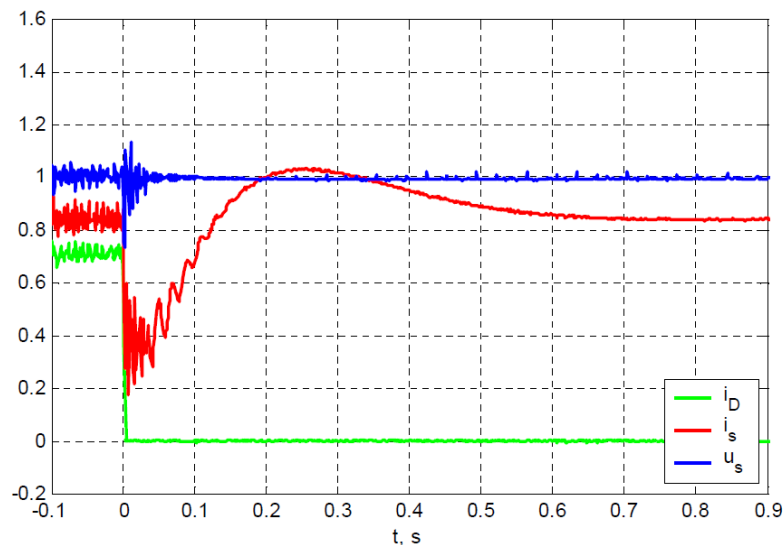


Рис. 2. Результаты моделирования переходных процессов пуска электропривода нагнетателя

Представлен благоприятный случай: к моменту окончания паузы вектор потокосцепления статора совпадает с вектором потокосцепления сети, совпадают также напряжения по модулю. В результате после подключения к сети свободная составляющая процесса практически отсутствует – колебания частоты 50 Гц в модуле вектора тока статора незначительны. Но имеют место переходные электромагнитные и электромеханические процессы: нарастание тока статора и восстановление скорости, – протекают с небольшим перерегулированием по току.

На рис. 3 представлены процессы при таких же условиях, как для рис. 2, но при уменьшенном начальном опережении вектора потокосцепления статора. На рис. 6

показаны процессы при таких же условиях, как для рис. 4, но при увеличенном напряжении сети.

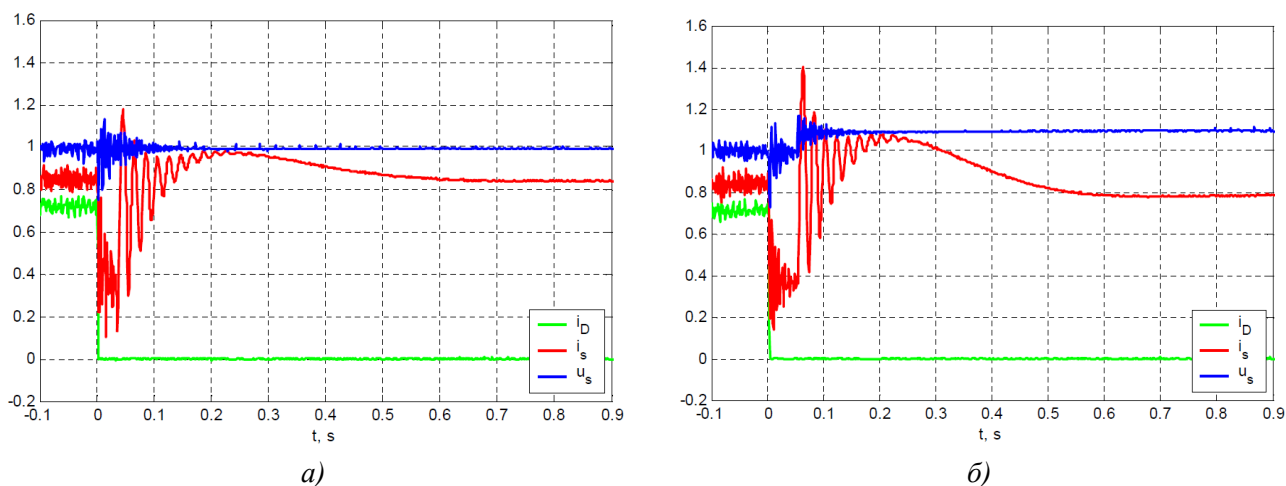


Рис. 4. Динамика процессов: а) при уменьшенном начальном опережении вектора потокосцепления статора, б) при увеличенном напряжении сети

В обоих случаях очень заметна свободная составляющая тока статора. Перерегулирование по току намного выше, чем в благоприятном случае. Следует отметить, что процессы заметно неблагоприятнее в малоинерционных приводах, например, в приводах циркуляционных насосов ТЭЦ.

Переключение с начальным опережением по скорости

В режим с опережением по скорости начальное опережение по скорости вводится для того, чтобы в момент подключения к сети угловая скорость вектора потока двигателя была близка к таковой для вектора потокосцепления сети. Необходимое опережение по скорости пропорционально моменту нагрузки двигателя. Желательно вводить опережение с учётом фактического момента нагрузки. В системе регулирования электропривода существует сигнал, пропорциональный нагрузке. Например, в электроприводе с ПИ регулятором скорости выходной сигнал интегральной части регулятора скорости пропорционален моменту нагрузки. На рис. 5 показаны процессы в системе регулирования со специальным регулятором скорости; это – процессы при скачкообразных изменениях момента нагрузки.

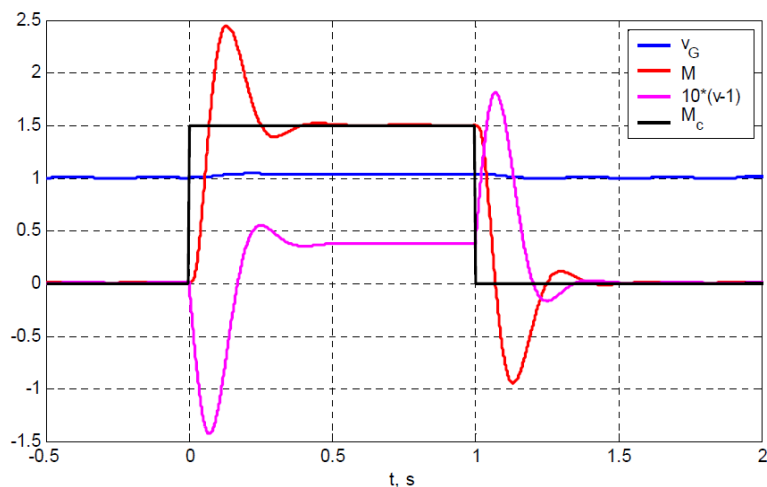


Рис. 5. Динамика процессов в системе регулирования со специальным регулятором скорости
Зарегистрированы: v_G – задание скорости, M – электромагнитный момент двигателя, $(v - 1)$ – отклонение скорости от скорости х. х., M_c – момент нагрузки. Посленаброса нагрузки происходит просадка скорости, затем устанавливается режим с

превышением задания скорости и фактической скорости над уровнем скорости х. х. Характер процессов практически не отличается от типовых процессов системы подчинённого регулирования с ПИ регулятором скорости. Таким образом, замкнутая система вполне устойчива.

Режимы с перекрытием включённого состояния контакторов $QF1$, $QB1$

В этом варианте вначале запирается ПРУ, затем включается контактор $QB1$ (статор подключается к сети), после чего отключается контактор $QF1$ (статор отключается от выхода ПРУ). Если время T_B включения контактора $QB1$ достаточно существенно, вначале должна подаваться команда на включение этого контактора. Команду на запираение ПРУ желательно подать с задержкой по времени, чтобы обеспечить малое время паузы T_P . Интервал времени T_B от подачи первой команды до подключения статора к сети содержит два подинтервала:

1. Подинтервал продолжительностью $(T_B - T_P)$, на котором вектор потока двигателя вращается с опережением по скорости $\Delta\gamma_0$.

2. Собственно пауза продолжительностью T_P , когда вектор потока двигателя тормозится. Соответственно, необходимое опережение по углу в момент подачи первой команды должно составлять:

$$\Delta\gamma_1 = \Omega_b \left(-T_B + \frac{3}{2} T_P \right) \frac{T_P}{T_j} M_c.$$

Режимы без перекрытия включённого состояния контакторов $QF1$, $QB1$

В этом случае возможны разные варианты в отношении последовательности команд. Рассматриваем самый неблагоприятный вариант в отношении времени паузы, но зато наиболее надёжный: гарантировано исключается перекрытие включённого состояния контакторов. В этом варианте контакторы $QF1$, $QB1$ взаимно блокированы, включение контактора $QB1$ только начинается после отключения контактора $QF1$. Чтобы исключить разрыв цепи статора при работающем ПРУ, команда на запираение ПРУ подаётся одновременно с командой на отключение статора от ПРУ (отключение контактора $QF1$).

В данном случае время паузы приближённо содержит время отключения контактора $QF1$ и время включения контактора $QB1$. Оно может оказаться на порядок больше, чем в варианте с перекрытием включённого состояния контакторов. Необходимое опережение по фазе в этом случае составляет:

$$\Delta\gamma_1 = \frac{1}{2} \Omega_b \frac{T_P^2}{T_j} M_c.$$

И опережение по скорости, и опережение по углу при значительном времени паузы оказываются приемлемыми только для высоко-инерционных электроприводов: вентиляторов, турбокомпрессоров и т.п. С 1997 г. эксплуатируется система частотного пуска синхронных двигателей 20 МВт турбокомпрессоров, разработанная под руководством авторов [21-23].

Выводы

Частотный пуск – высокоэффективный метод пуска двигателей переменного тока, не нуждающихся в регулировании скорости. Автоматическая синхронизация с питающей сетью и автоматическое безударное переключение статора двигателя на сеть – достаточно сложный, ответственный и, в определённой мере, рискованный режим. Гладкое переключение на сеть обеспечивается только в варианте переключения на сеть без паузы.

Библиографический список

1. Sonin Y.P., Bainev V.F., Gulyaev I.V. Static characteristics of contactless asynhronizedthyatron motor // Russian Electrical Engineering. 1994. Т. 65. № 9. С. 20-29.
2. Kopylov I.P., Sonin Yu.P., Gulyaev I.V., Nikulin V.V. Contactless asynhronized synchronous motor // Russian Electrical Engineering. 1999. Т. 70. № 9. С. 35-41.
3. Крюков О.В., Мещеряков В.Н., Гуляев И.В. Электроприводы на основе машины двойного питания и АВК с преобразователями в цепях статора и ротора // Саранск, 2020.
4. Belousov A.S., Meshcheryakov V.N., Valtchev S., Kryukov O.V. Development of a control algorithm for three-phase inverter in two-phase electric drives reducing the number of commutations // В сборнике: Proceedings - 2019 1st International Conference on Control Systems, Mathematical Modelling, Automation and Energy Efficiency, SUMMA 2019. 2019. С. 444-449.
5. Крюков О.В. Мониторинг условий эксплуатации электродвигателей газоперекачивающих агрегатов // Контроль. Диагностика. 2016. № 12. С. 50-58.
6. Васенин А.Б., Крюков О.В. Оценка эксплуатационных факторов электроприводных ГПА // Электрооборудование: эксплуатация и ремонт. 2019. № 7. С. 6-15.
7. Крюков О.В., Степанов С.Е. Модернизация систем управления ЭГПА в условиях действующих компрессорных станций // В сборнике: Проблемы автоматизации и управления в технических системах. МНТК под редакцией М.А. Щербакова. 2013. С. 29-32.
8. Крюков О.В., Степанов С.Е., Серебряков А.В. Современный подход к организации ремонта по данным прогноза технического состояния и ресурса электрооборудования // Газовая промышленность. 2017. № 8 (756). С. 84-89.
9. Васенин А.Б., Крюков О.В. Проектирование электромеханической части и систем управления энергетических установок газотранспортных потребителей // Известия Тульского гос. ун-та. Технические науки. 2011. № 5-1. С. 47-51.
10. Васенин А.Б., Степанов С.Е., Крюков О.В. Методология и средства оперативного мониторинга электродвигателей на КС // Контроль. Диагностика. 2019. № 11. С. 52-58.
11. Сулин В.А., Крюков О.В. Особенности развития метрологических расходоизмерительных центров // Наука и техника в газовой промышленности. 2020. № 2 (82). С. 99-106.
12. Крюков О.В., Туганов Р.Б., Мещеряков В.Н. Компьютерное моделирование синхронных машин при эксплуатации на объектах энергетики // Автоматизация и ИТ в энергетике. 2020. № 4 (129). С. 26-31.
13. Ilin M.V., Bespalov N.N., Kapitonov S.S., Gulyaev I.V. Selection of power semiconductor devices for parallel connection // Russian Electrical Engineering. 2017. Т. 88. № 6. С. 336-341.
14. Kryukov O.V., Gulyaev I.V., Teplukhov D.Y. Method for stabilizing the operation of synchronous machines using a virtual load sensor // Russian Electrical Engineering. 2019. Т. 90. № 7. С. 473-478.
15. Мещеряков В.Н., Ласточкин Д.В., Крюков О.В. Приложения теории нечетких множеств для обработки данных и задач прогнозирования в системах АЭП // В сборнике: Современные сложные системы управления. Материалы XII МНПК. 2017. С. 153-158.
16. Крюков О.В., Серебряков А.В. Активно-адаптивные алгоритмы управления и мониторинга автономными энергетическими комплексами // В сборнике: Пром-Инжиниринг. Труды II МНТК. ЮУрГУ. 2016. С. 286-290.
17. Vasenin A.B., Kryukov O.V., Serebryakov A.V. Adaptive control algorithms of autonomous generator complexes // В книге: Электромеханика, электротехнологии, электротехнические материалы и компоненты. Труды МКЭЭЭ-2016. 2016. С. 133-135.

18. Крюков О.В. Алгоритмы быстрого преобразования Уолша в микропроцессорных системах управления электроприводом // Изв. ВУЗов. Электромеханика. 2005. № 4. С. 39-44.
19. Крюков О.В., Краснов Д.В. Перспективы применения преобразователей частоты для регулирования производительности ЭГПА // Газовая промышленность. 2014. № 6. С. 86-89.
20. Kryukov O.V., Blagodarov D.A., Dulnev N.N., Safonov Y.M., Fedortsov N.N., Kostin A.A. Intelligent control of electric machine drive systems // В сборнике: 2018 10th International Conference on Electrical Power Drive Systems, ICEPDS 2018 - Conference Proceedings 10. 2018.
21. Васенин А.Б., Крюков О.В. Энергоэффективные и экологичные установки воздушного охлаждения // В сборнике: Великие реки' 2017 труды научного конгресса 19-го МНПФ. НГАСУ. 2017. С. 93-96.
22. Крюков О.В. Комплексный анализ условий эксплуатации электродвигателей ГПА // Компрессорная техника и пневматика. 2013. № 4. С. 14-19.
23. Мещеряков В.Н., Туганов Р.Б., Крюков О.В. Способы интеллектуального управления электроприводными агрегатами газопровода "Парабель-Кузбасс" // Автоматизация и ИТ в нефтегазовой области. 2020. № 1 (39). С. 5-10.

A.B. Vasenin, I.V. Gulyaev, O.V. Kryukov

DYNAMICS OF THE FREQUENCY START OF AUTOMATED ELECTRIC DRIVE

*Gazprom design LLC, Nizhny Novgorod,
Mordovia State University named N. P. Ogarev, Saransk,
TSN-electro LLC, Nizhny Novgorod*

High-efficiency methods of frequency starting of engines of the main technological units of the fuel and energy complex are considered. The analysis of circuit design options for frequency start-up showed the possibility of significant improvement of all the main disadvantages of direct start-up. The advantages of automatic synchronization with the supply network and automatic shock-free switching of the motor stator to the network are presented. Best practices of smooth switching to the network is provided only in the version with the algorithm of operation without pause when powering the stator circuit with a minimum of over-regulation of the stator current and fluctuations in the electromagnetic torque of the motor.

Keywords: frequency start, electric motor, starting control device, pause, switching, simulation.

СРЕДСТВА НАДЕЖНОГО ЭНЕРГООБЕСПЕЧЕНИЯ ОБОРУДОВАНИЯ ПОДВОДНЫХ ДОБЫЧНЫХ КОМПЛЕКСОВ

ООО «ТСН-электро», г. Нижний Новгород, ГУМРФ им. адм. С.О. Макарова,
г. Санкт-Петербург, ВГУВТ, г. Нижний Новгород

Рассмотрена стратегия развития систем энергообеспечения подводных добычных комплексов исходя из специфических особенностей акваторий Арктики РФ. Показана возможность реализации проектов морских месторождений углеводородного сырья на основе новой аппаратной базы электромашиностроения и интеллектуальных алгоритмов управления и мониторинга электроприводными газоперекачивающими агрегатами. Представлены высоконадежные системы электроснабжения, которые обеспечивают создание высокотехнологичных, энергоэффективных и долговечных агрегатов и автоматизированных систем на основе безлюдных технологий. Рассмотрены особенности реализации новых оригинальных технико-технологических решений функциональных возможностей электроприводных газоперекачивающих агрегатов с интеллектуальными системами управления и диагностики.

Ключевые слова: системы электроснабжения, подводные добычные комплексы, электроприводной газоперекачивающий агрегат, интеллектуальные системы диагностики, безлюдные технологии.

Введение

В настоящее время потребление углеводородного сырья в мире возрастает из года в год, месторождения на суше истощаются, а открытие новых происходит в основном в шельфовой зоне материков или в глубоководной части океана. Обустройство таких месторождений требует нового подхода, поскольку шельфовые месторождения РФ (рис. 1) расположены в основном в акватории замерзающих морей Арктики и Дальнего Востока.



Рис. 1. Шельфовая и морская добыча природного газа в РФ

Одной из проблем при обустройстве морских удаленных месторождений нефти и газа является сложность обеспечения надежного энергоснабжения [1-3]. Из-за большого расстояния от берега в силовом кабеле возникают потери, обусловленные использованием систем переменного тока, который нужен для электроприводов технологических агрегатов [4-7]. Существующие системы электроснабжения подводно-добычных комплексов (ПДК) нефтегазовых месторождений в основном реализованы на шельфе Норвегии. Изучены принципы подводной передачи энергии на большие расстояния с помощью высоковольтных кабелей постоянного тока и особенности применения

автономных инверторов и оригинальных конструкций приводных электродвигателей в системе технологических процессов компримирования на ПДК [8-10].

Ввиду большого количества необустроенных месторождений газа в Арктическом регионе России и короткого навигационного периода тема данной работы является актуальной, так как разработка шельфовых месторождений в замерзающих морях возможна только подводным способом. Развитие системы энергоснабжения ПДК позволит в короткие сроки ввести в эксплуатацию труднодоступные месторождения в акваториях Тихого и Северного Ледовитого океана.

В настоящее время электроприводные нефте- и газоперекачивающие агрегаты являются единственным и безальтернативным вариантом реализации энергетических систем обустройства на подводных добычных комплексах (ПДК) [11-13]. Новая аппаратная база электромашиностроения и интеллектуальные алгоритмы управления и мониторинга [14-17] обеспечивают создание высокотехнологичных, энергоэффективных и надежных агрегатов и автоматизированных систем на основе безлюдных технологий [18-21]. Особенно это актуально в сочетании с комплексом систем энергообеспечения промышленных объектов, где высоки требования по живучести объектов обустройства подводных месторождений.

Рассмотрим особенности реализации новых оригинальных технико-технологических решений функциональных возможностей электроприводных газоперекачивающих агрегатов (ЭГПА) с интеллектуальными системами управления и диагностики (ИСУД): безредукторное и безмасляное исполнение в едином монокорпусе с нагнетателями; надежный электромагнитный подвес ротора; плавный безопасный пуск под нагрузкой; частотное и частотно-токовое релейное регулирование параметров и автоматическая адаптивная перестройка законов управления; встроенная система мониторинга и прогнозирования технического состояния.

Обустройство подводных добычных комплексов

Сегодня на мировом рынке среди компаний проектантов и изготовителей подводного оборудования мировое лидерство у следующих компаний: FMC KongsbergSubsea AS, AkerSolutions (Subsea), Cameron и GE Vetco. Подводное устьевое оборудование представляет собой комплекс специальных механизмов, устройств и систем, обеспечивающих при бурении разведочных скважин и добычи в штатном режиме механическую связь буровой установки с устьем скважины, расположенном на дне моря.

Технологии подводной подготовки углеводородов существенно расширяют гибкость в добыче продукции скважин. В состав подводного комплекса входит: центробежный газовый компрессор; электропривод с системой охлаждения; оборудование регулирования частоты вращения привода компрессора; оборудование подачи и распределения электропитания для потребителей блока; дистанционно-управляемая запорная арматура; контрольно-измерительные приборы; системы управления, аварийного выключения, мониторинга текущего состояния.

Несомненным лидером в области применения подводных технологий при освоении шельфовых месторождений нефти и газа является Норвегия. Она сумела одновременно создать свою национальную инновационную систему и сделать ее частью глобальной, добиться, чтобы иностранные корпорации, проводили локализацию своих технологий в стране.

Еще в 2001 году в Норвегии была принята национальная стратегия «Нефть и газ в 21 веке» (OG21), которая определила 8 направлений исследований, а в 2008 году Научно-исследовательскому совету выделено 5691 миллион крон (более 28,5 млрд рублей). Один из проектов - WS SeabedRig DEMO2000 по испытаниям полностью автоматизированного ПДК (рис. 2).

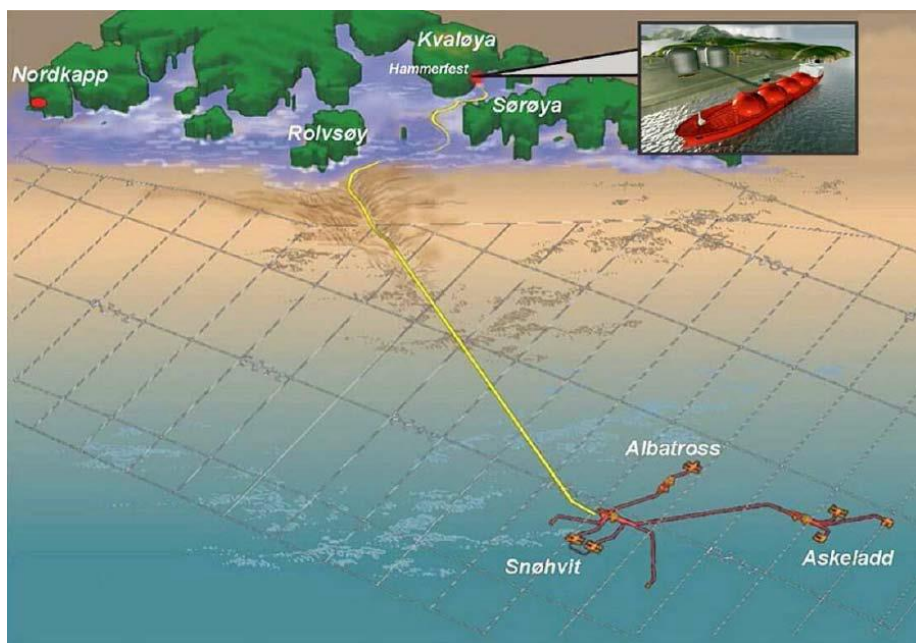


Рис. 2. Схема реализации проекта, автоматизированного ПДК

Технологии атомной энергетики для обустройства ПДК

Учитывая удельные капитальные вложения и прогнозируемую себестоимость электроэнергии, можно говорить о целесообразности и экономической эффективности атомных энергоисточников единичной мощности, не превышающей 300 МВт (что по определению МАГАТЭ соответствует диапазону станций малых мощностей (АСММ)), при решении проблемы энергообеспечения перспективных арктических месторождений.

Главное требование к энергетическим установкам в условиях Арктического шельфа – повышенная надежность и минимальное воздействие на окружающую среду. Не менее актуально требование минимального обслуживания вплоть до полной автономности с использованием безлюдных технологий. Таким условиям идеально отвечают судовые атомные энергетические установки (рис. 3).

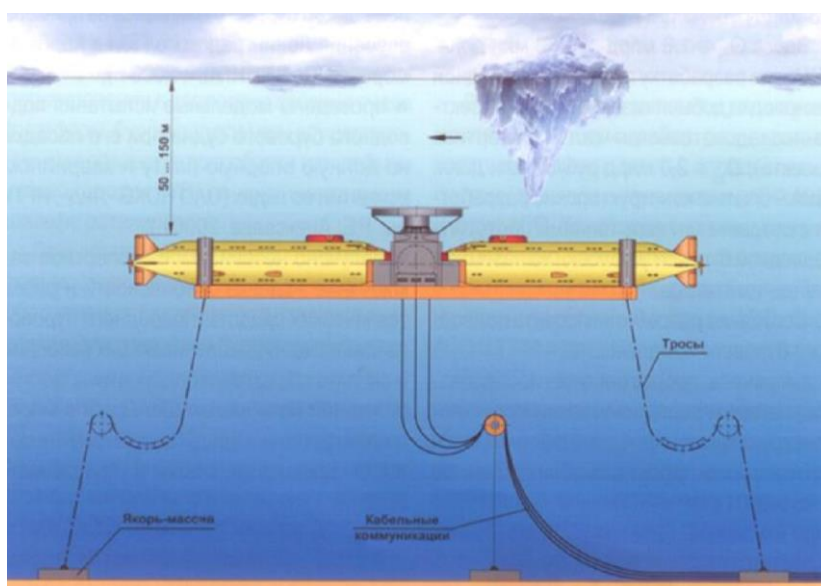


Рис. 3. Атомная энергоустановка на базе реакторной установки «Шельф»

Интегральный опыт эксплуатации малой атомной транспортной энергетики превышает шесть тысяч реакторо-лет. Промышленностью произведено более 500 реакторных установок.

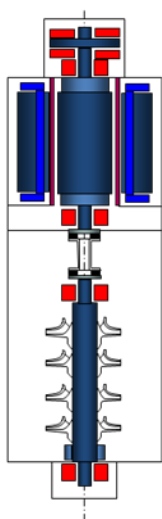
Так, например, предлагается атомная турбогенераторная установка (АТГУ) «Шельф» для энергоснабжения технических средств, работающих на нефтегазовых месторождениях, в том числе удаленных на значительное расстояние от берега и имеющих круглогодичный цикл работы в течение 25–30 лет. Она может поставляться в виде готовой к эксплуатации энергокапсулы в составе наземного или подводного исполнения.

АТГУ «Шельф» включает двухконтурную атомную РУ с водо-водяным интегральным реактором тепловой мощностью 28 МВт, турбогенераторную установку, обеспечивающую выработку электроэнергии мощностью 6000 кВт и систему автоматизированного и дистанционного управления, контроля и защиты техническими средствами установки.

Особенности реализации электроприводных газоперекачивающих агрегатов

Ужесточение требований по реализации новых оригинальных технико-технологических решений функциональных возможностей ЭГПА с интеллектуальными системами управления и диагностики обусловили создание новой техники: безредукторное и безмасляное исполнение в едином монокорпусе с нагнетателями; надежный электромагнитный подвес ротора; плавный безопасный пуск под нагрузкой; частотное и частотно-токовое релейное регулирование параметров и автоматическая адаптивная перестройка законов управления; встроенная система мониторинга и прогнозирования технического состояния.

Повышение требований к минимизации массогабаритных показателей КС для компактности «пятна застройки», обеспечения безопасности для окружающей среды с «нулевыми выбросами» при работе с грязными газами без внешней спецочистки и увеличение надежности и долговечности работы установок предложен вариант вертикальной компоновки приводных двигателей ЭГПА в капсулированном герметичном корпусе без мультипликатора, муфты и маслосистем (рис. 4).



- Осевой магнитный подшипник в герметичном корпусе
- Радиальный магнитный подшипник в герметичном корпусе
- Массивный ротор двигателя
- Асинхронный двигатель в герметичной оболочке
- Общий защитный кожух под давлением
- Муфта трубчатого вала
- Многоступенчатый центробежный компрессор
- Цельный ротор
- Радиальный магнитный подшипник в герметичном корпусе

Рис. 4. Схема вертикальной компоновки ЭГПА.

Конструктивно ЭГПА имеет компактную обмотку статора с жидкостным охлаждением и массивный ротор. Технология с компактными торцевыми обмотками применялась ранее для тихоходных АД специального назначения впервые применена для высокоскоростных электроприводов.

Капсулированные ЭГПА с вертикальной осью вращения имеют следующие параметры:

- номинальные мощности 7,5; 9,5; 12,2 и 15 МВт,
- номинальные скорости 7600 - 20000 об/мин,

- количество рабочих колес компрессора 6 и 8,
- схема колес – последовательная и сдвоенная,
- капсула: длина 4,8 м, ширина 4,1 м, высота 4,6 м,
- снижение пятна застройки – более 40 %,
- герметичная оболочка держит 150 бар и 200 °С.

Главным направлением дальнейшего совершенствования мощных электрических машин переменного тока для нужд ПДК сегодня является повышение надежности, ресурса и энергоэффективности установок. Это связано с постоянным ужесточением требований по их эксплуатации на объектах повышенной опасности (ОПО), с одной стороны, и тенденциями интеграции приводных электрических машин с исполнительными механизмами и системами автоматического управления, с другой.

Повышение требований к минимизации массогабаритных показателей ПДК для компактности «пятна застройки», обеспечения безопасности для окружающей среды с «нулевыми выбросами» при работе с грязными газами без внешней спецочистки и увеличение надежности и долговечности работы установок предложен вариант вертикальной компоновки приводных двигателей ЭГПА в капсулированном герметичном корпусе без мультипликатора, муфты и маслосистем.

Подобные уникальные конструктивные решения позволяют реализовать малолюдные и безлюдные технологии эксплуатации компрессорных установок с автоматическим дистанционным управлением и оптимизацией работы всего подводного добычного комплекса. Кроме того, имеются практические наработки характерных примеров многолетней успешной промышленной эксплуатации инновационных систем электроснабжения с секционирующими пунктами на базе реклоузеров для электроприводных газо- и нефтеперекачивающих агрегатов, обеспечивающих реализацию малолюдных и безлюдных технологий, включая автономные системы энергообеспечения арктических месторождений углеводородов. Предложены модульные инвариантные структуры и законы оптимального управления системами энергообеспечения потребителей обустройства морских месторождений нефти и газа, которые позволяют гарантировать энергоэффективное электропитание основных технологических установок и других потребителей подводного добычного комплекса.

Выводы

Новые конструктивные решения систем энергообеспечения и электроснабжения ПДК на базе электродвигателей переменного тока мегаваттного класса с горизонтальной и вертикальной осями вращения, а также совершенствование систем управления частотно-регулируемым электроприводом, позволяют повысить надежность и энергоэффективность технологических агрегатов в рамках безлюдных технологий для важнейшей топливно-энергетической отрасли промышленности.

Библиографический список

1. Захаров П.А., Титов В.Г., Крюков О.В. Информационные технологии реализации энергообеспечения ПДК // В сборнике: Информационные системы и технологии ИСТ-2020. XXVI МНТК. НГТУ им. Р.Е. Алексеева. 2020. С. 300-306.
2. Туганов Р.Б., Крюков О.В. Атомные энергоустановки в энергообеспечении подводных добычных комплексов // В сб.: Великие реки' 2019. труды научного конгресса. 2019. С. 47-50.
3. Ковтун Л.И., Крюков О.В., Саушев А.В., Антоненко С.И. Аналитико-статистический метод оценки состояния и прогнозирования рисков сложных технических систем // Труды Международного симпозиума «Надежность и качество». Пенза. 2020. Т.1. С. 264-269.

4. Kryukov O.V., Gulyaev I.V., Teplukhov D.Y. Method for stabilizing the operation of synchronous machines using a virtual load sensor // Russian Electrical Engineering. 2019. T. 90. № 7. С. 473-478.
5. Vasenin A.B., Kryukov O.V., Serebryakov A.V. Adaptive control algorithms of autonomous generator complexes // В книге: Электромеханика, электротехнологии, электротехнические материалы и компоненты. Труды МКЭЭЭ-2016. 2016. С. 133-135.
6. Васенин А.Б., Крюков О.В. Проектирование электромеханической части и систем управления энергетических установок газотранспортных потребителей // Известия Тульского гос. ун-та. Технические науки. 2011. № 5-1. С. 47-51.
7. Крюков О.В., Степанов С.Е. Модернизация систем управления ЭГПА в условиях действующих компрессорных станций // В сборнике: Проблемы автоматизации и управления в технических системах. МНТК под редакцией М.А. Щербакова. 2013. С. 29-32.
8. Мещеряков В.Н., Крюков О.В. Научные основы и энергосберегающие технологии обеспечения ПДК // В книге: Газотранспортные системы: настоящее и будущее (GTS-2019). Тезисы докладов по материалам VIII МНТК. 2019. С. 41.
9. Belousov A.S., Meshcheryakov V.N., Valtchev S., Kryukov O.V. Development of a control algorithm for three-phase inverter in two-phase electric drives reducing the number of commutations // В сборнике: Proceedings - 2019 1st International Conference on Control Systems, Mathematical Modelling, Automation and Energy Efficiency, SUMMA 2019. 2019. С. 444-449.
10. Васенин А.Б., Степанов С.Е., Крюков О.В. Методология и средства оперативного мониторинга электродвигателей на КС // Контроль. Диагностика. 2019. № 11. С. 52-58.
11. Крюков О.В. Энергоэффективные и экологичные электростанции для удаленных промышленных потребителей // Экологические системы и приборы. 2018. № 3. С. 18-25.
12. Крюков О.В., Мещеряков В.Н., Гуляев И.В. Электроприводы на основе машины двойного питания и асинхронного вентильного каскада с преобразователями в цепях статора и ротора // Саранск, 2020.
13. Крюков О.В. Особенности энергоснабжения подводных комплексов морских месторождений углеводородов // В сборнике: Фёдоровские чтения-2019. Под общей ред. Б.И. Кудрина, Ю.В. Матюниной. 2019. С. 357-359.
14. Мещеряков В.Н., Ласточкин Д.В., Крюков О.В. Приложения теории нечетких множеств для обработки данных и задач прогнозирования в системах АЭП // В сборнике: Современные сложные системы управления. Материалы XII МНПК. 2017. С. 153-158.
15. Крюков О.В., Саушев А.В. Методология вейвлет-преобразования при анализе автоматизированных электроприводов // В сб.: Фёдоровские чтения - 2020: L МНПК под общ. ред. Ю.В. Матюниной. М.: ИД «МЭИ». 2020. С. 174-178.
16. Крюков О.В. Алгоритмы быстрого преобразования Уолша в микропроцессорных системах управления электроприводом // Известия ВУЗов. Электромеханика. 2005. № 4. С. 39-44.
17. Васенин А.Б., Крюков О.В. Энергоэффективные и экологичные установки воздушного охлаждения // В сборнике: ВЕЛИКИЕ РЕКИ' 2017 труды научного конгресса 19-го МНПФ. НГАСУ. 2017. С. 93-96.
18. Крюков О.В. Мониторинг условий эксплуатации электродвигателей газоперекачивающих агрегатов // Контроль. Диагностика. 2016. № 12. С. 50-58.
19. Крюков О.В., Степанов С.Е., Серебряков А.В. Современный подход к организации ремонта по данным прогноза технического состояния и ресурса электрооборудования // Газовая промышленность. 2017. № 8 (756). С. 84-89.

20. Крюков О.В., Серебряков А.В. Активно-адаптивные алгоритмы управления и мониторинга автономными энергетическими комплексами // В сборнике: Пром-Инжиниринг. Труды II МНТК. ЮУрГУ. 2016. С. 286-290.
21. Крюков О.В. Комплексный анализ условий эксплуатации электродвигателей ГПА // Компрессорная техника и пневматика. 2013. № 4. С. 14-19.

O. V. Kryukov, A. V. Saushev, R. B. Tuganov

**MEANS OF RELIABLE POWER SUPPLY OF EQUIPMENT OF UNDERWATER
MINING COMPLEXES**

*LLC "TSN-electro", Nizhny Novgorod, GUMRF named after adm. S. O. Makarov,
St. Petersburg, VGUVT, Nizhny Novgorod*

The strategy of development of energy supply systems for subsea mining complexes based on the specific features of the Arctic waters of the Russian Federation is considered. The possibility of implementing projects of offshore hydrocarbon deposits on the basis of a new hardware base of electrical engineering and intelligent algorithms for controlling and monitoring electric gas pumping units is shown. Highly reliable power supply systems are presented, which ensure the creation of high-tech, energy-efficient and durable units and automated systems based on unpopulated technologies. The features of the implementation of new original technical and technological solutions of the functional capabilities of electric gas pumping units with intelligent control and diagnostic systems are considered.

Keywords: power supply systems, underwater mining complexes, electric gas pumping unit, intelligent diagnostic systems, unpopulated technologies.

ЦИФРОВИЗАЦИЯ ПОДСТАНЦИИ ГПП-10 ПАО НЛМК*Липецкий государственный технический университет*

«Цифровая подстанция» позволяет получить единый цифровой поток данных, характеризующий состояние управляемого объекта. Абстрагирование от существующей парадигмы построения системы защиты и управления подстанцией и переход к программной платформе, размещаемой на универсальных аппаратных устройствах и имеющей свободное распределение функций дает возможность получать решения, имеющие как полностью распределённую, так и централизованную архитектуру. Применение единой программной платформы позволит видоизменить рынок аппаратных устройств для построения систем защиты и управления подстанцией и перейти к рынку функциональных алгоритмов.

Ключевые слова: цифровая подстанция, релейная защита и автоматика, АСУТП ПС, SCADA, электроснабжение.

Цифровая подстанция – подстанция, оборудованная комплексом цифровых устройств (терминалов) для решения задач контроля и учета качества электроэнергии, регистрации аварийных событий, телемеханики, релейной защиты и автоматики (РЗА) и АСУТП. Передача информации между оборудованием объединяется центральным сервером объекта по последовательным каналам связи на единых протоколах.

Передача сигналов в цифровом виде на всех уровнях управления подстанцией позволит создать технологическую инфраструктуру для внедрения информационно-аналитических систем, снизить ошибки недоучета электроэнергии, уменьшить затраты на обслуживание подстанции, а также повысить электромагнитную безопасность и надежность работы микропроцессорных устройств. Внедрение систем, удовлетворяющих стандарту МЭК 61850 «Сети и системы связи на подстанциях», обеспечит высокую безопасность и скорость передачи информации, взаимозаменяемость отдельных элементов системы, повышение надежности системы.

Основные элементы, которые позволили спроектировать подобное решение, базируются на собственных разработках двух компаний: цифровые оптические измерительные трансформаторы компании «Профотек» и цифровая система защиты и управления компании «ЛИСИС». Это уникальный альянс российских компаний, которые предлагают эффективное решение, не имеющее сегодня аналогов в мире.

Группа НЛМК – лидирующий международный производитель высококачественной стальной продукции с вертикально-интегрированной моделью бизнеса. Добыча сырья и производство стали сосредоточены в регионах, изготовление готовой продукции осуществляется в непосредственной близости от основных потребителей в России, Северной Америке и странах ЕС.

Благодаря самообеспеченности в основных сырьевых материалах, энергии и высокой технологической оснащенности НЛМК входит в число самых эффективных и прибыльных производителей стали в мире. Компания имеет диверсифицированный продуктовый портфель, обеспечивающий лидерство на локальных рынках и высокую эффективность продаж. Используя преимущества гибкой производственной цепочки, сбалансированного продуктового ряда, эффективной системы сбыта и обширной географии продаж, Компания обладает возможностью своевременно реагировать на постоянно изменяющиеся рыночные условия.

Микропроцессорные реле позволяют создать автоматизированную систему управления технологическим процессом подстанции. При включении в систему SCADA появляется возможность контроля за режимом работы подстанции с удаленного рабочего места, получение информации о состоянии оборудования в реальном времени, показателей качества электрической энергии. Также появляется возможность производства оперативных переключений дистанционно с рабочего места диспетчера, а также возможность контроля выполнения переключений подчиненным оперативным персоналом. Цель модернизации – замена электромеханических реле цифровыми терминалами защиты.

Подстанция ГПП-10 является подстанцией глубокого ввода, с 4-мя трансформаторами по 63000 кВА, сооружена для питания потребителей электроэнергии производства холодного проката и покрытий. Питание ГПП-10 осуществляется на напряжение 110кВ по четырем кабельным линиям от распределительной подстанции 110кВ РП-1. Питание РП-1 осуществляется двумя линиями 110 кВ от подстанции «Северная» и двумя линиями 110 кВ от подстанции «Новая». Схема внешнего электроснабжения производства холодного проката и покрытий (ПХПП) приведена на рисунке 1.

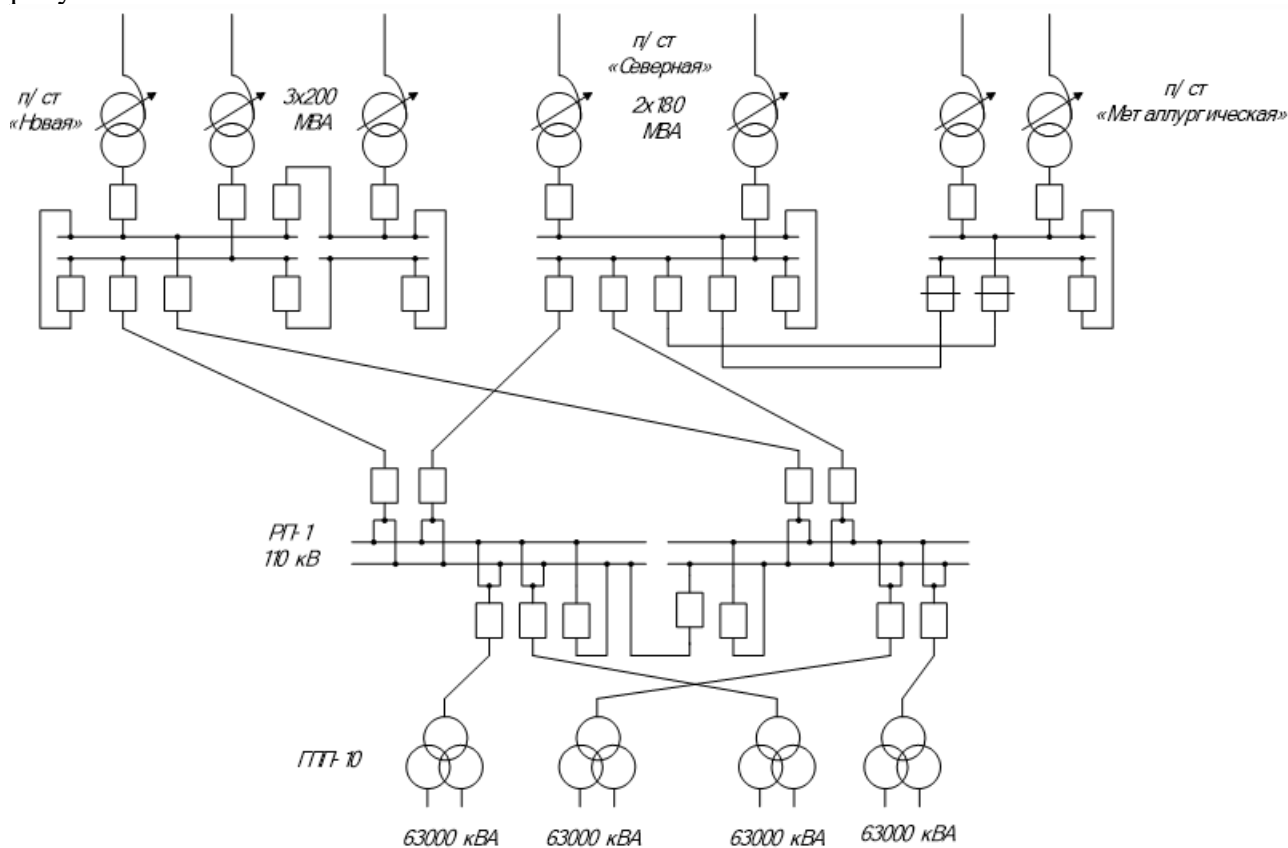


Рис. 1. Схема внешнего электроснабжения ПХПП

В связи со стесненностью площадки в районе цеха питание ГПП-10 осуществляется радиальными кабельными линиями 110 кВ по схеме «Кабельная линия-трансформатор с кабельными вводами 110кВ ТРДЦНК 110/10/10 63МВА-токопровод 10кВ», с передачей отключающего импульса по кабелю связи на РП1-110кВ.

Переход на новую элементную базу не приводит к изменению принципов релейной защиты и автоматики, а только расширяет её функциональные возможности, упрощает эксплуатацию и повышают точность. По этим причинам микропроцессорные реле быстро занимают место электромеханических.

В качестве микропроцессорной релейной защиты ГПП-10 предлагается выбрать серию микропроцессорных терминалов РЗА MiCOM фирм Alstom Grid и Schneider Electric.

Большинство реле серий MiCOM 40 могут поставляться с интерфейсами Ethernet, что позволяет реализовать полное решение для подстанции в соответствии с МЭК 61850 [1].

В зону действия дифференциальной защиты трансформаторов 1Т и 2Т входит реактор РБДГ10-4000-0,105. Вместо имеющейся дифференциальной защиты принимается защита на базе устройства MiCOM P643. Максимальная токовая защита на стороне 110 кВ защищаемого трансформатора для резервирования основных защит трансформатора и от сверхтоков внешних коротких замыканий. Защита выполнена с комбинированным пуском по напряжению со стороны 10 кВ. Принимается защита на базе устройства MiCOM P143. В соответствии с ПУЭ [2] на вводах 10 кВ трансформаторов предусматривается устройство автоматического повторного включения – АПВ с проверкой отсутствия напряжения. Пуск АПВ осуществляется от максимальной токовой защиты, установленной на вводах 10 кВ от трансформаторов. При к.з. на шинах 10 кВ, реле пуска АПВ осуществляется запрет действия АВР. Для выполнения функции АПВ предлагается использовать устройство MiCOM P145. Для выполнения функции АВР, контроля цепей напряжения и контроля изоляции сети принимается устройство MiCOM P922. Также на подстанции имеется автоматическая частотная разгрузка на базе устройства АЧР типа MiCOM P143 с контролем скорости синхронизма. Характеристики реле представлены в таблице 1.

Таблица 1

Характеристики MiCOM 40 для ГПП

Тип реле	MiCOM P643	MiCOM P143	MiCOM P145
Функционал	Дифференциальная защита трансформатора	Защита от межфазных КЗ (МТЗ) с обеспечением резервирования отключения междуфазных КЗ.	Направленная трехфазная МТЗ и защита по напряжению, многофункциональная
Наличие IEC61850	Да	Да	Да
RS232	Да	Да	Да
Прочее	Дискретные входы, релейные выходы, выходы с высокой коммутационной способностью, оптоканалы	Дискретные входы, релейные выходы, выходы с высокой коммутационной способностью, оптоканалы	Дискретные входы, релейные выходы, выходы с высокой коммутационной способностью, оптоканалы

Устройства релейной защиты MiCOM в отношении электромагнитной совместимости отвечают стандартам EN Европейского Комитета по стандартизации в области электротехники. С точки зрения требований к помехозащищенности не все микропроцессорные устройства защиты соответствуют требованиям по степени жесткости испытаний к отдельным видам помех с сохранением качества функционирования, установленным в России.

Под автоматизированной системой управления технологическими процессами подстанции (АСУ ТП ПС) понимается комплекс аппаратных и программных средств для обеспечения контроля и управления процессом распределения электрической энергии, регистрации и предотвращения аварийных ситуаций, сбора данных о процессе и представления их в удобном для человека виде. Система интегрирует в себя информацию со всех смежных подстанционных систем, и передает данную информацию на верхние уровни иерархии управления режимами и эксплуатацией электрических сетей. АСУ ТП ПС обеспечивает возможность удаленного управления объектами и контроля параметров систем подстанции. Коммуникации внутри системы осуществляются с использованием стандартных промышленных протоколов, таких как МЭК 60870-5-10x, МЭК 61850, DNP 3.0, Modbus, Profibus.

Построение автоматизированной системы контроля и учета электроэнергии производится на базе многофункционального счетчика электроэнергии «BINOM337» производства ООО «Энфорс» [3], объединяющего в себе функции счетчика электроэнергии, прибора для измерения показателей качества электроэнергии, устройства

сбора и передачи данных. Это позволяет с помощью нескольких устройств полностью охватить информационную емкость КРУ по коммерческому и техническому учету.

Общая схема структуры АСУ ТП ГПП-10 представлена на рис. 2.

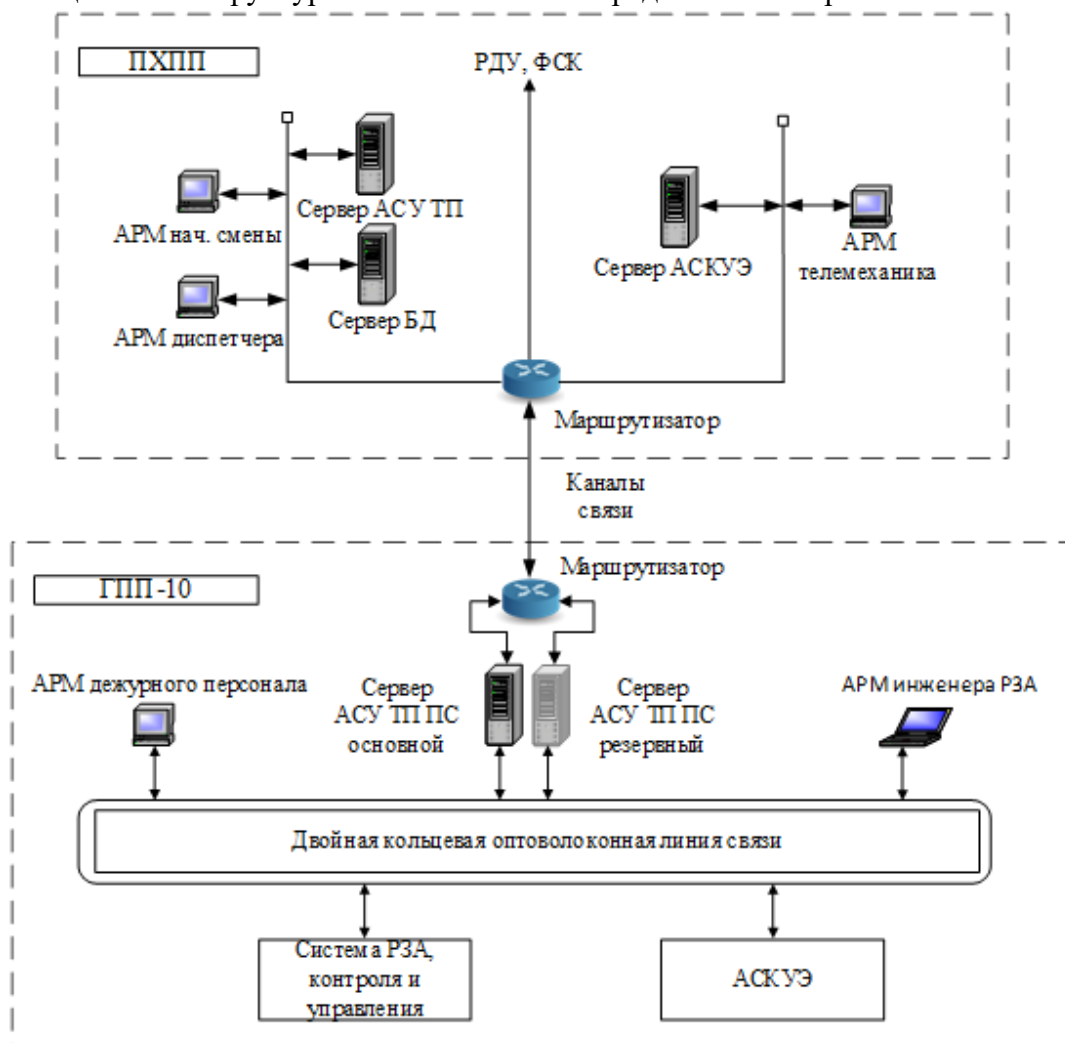


Рис. 2. Общая схема структуры АСУ ТП ПС ГПП-10

Для установки на вводные ячейки принимаются счетчики BINOM337, для отходящих присоединений и секционных выключателей – СЭТ-4ТМ.02.2. Соединение СЭТ-4ТМ.02.2 со счетчиками BINOM337 производится по интерфейсу RS-485. Для каждой секции счетчик, установленный на ячейке вводного выключателя, является одновременно концентратором данных.

Технические характеристики BINOM337 и СЭТ-4ТМ.02 сведены в таблицу 2.

Таблица 2

Технические характеристики счетчиков

Характеристика	BINOM337	СЭТ-4ТМ.02
Диапазон рабочих токов, $I_{ном}$	0,01...1,2	0,01...1,5
Класс точности активная/реактивная энергия	0,2S/0,5	0,2S/0,5
Интерфейс RS 485	Да	Да
Интерфейс RS 232	Да	Нет
Интерфейс Ethernet	Да	Нет
Оптический порт	Да	Да
Функции сбора и передачи данных	Да	Нет
Измерение показателей качества энергии	Да	Нет

Наиболее важными параметрами интерфейса являются пропускная способность и максимальная длина подключаемого кабеля. Промышленные интерфейсы обычно обеспечивают гальваническую развязку между соединяемыми устройствами. Наиболее распространены в промышленной автоматизации последовательные интерфейсы RS-485, RS-232, RS-422, Ethernet, CAN, HART, AS-интерфейс [4].

В качестве используемых интерфейсов связи на всех уровнях АСУ ТП ПС предполагается использование интерфейсов RS-485 (для АСКУЭ) и Ethernet.

Основными протоколами передачи данных, согласно стандарту МЭК-61850, являются протоколы MMS и GOOSE. MMS используется для передачи данных от терминалов РЗА в SCADA систему для дальнейшей визуализации, а GOOSE — для обмена данными между терминалами [5-6].

Для создания оптоволоконных колец и подключения к ним терминалов РЗА, а также для обеспечения связи системы АСКУЭ с кольцами применяются промышленные коммутаторы производителя Advantech – EKI-7248G-4X.

Так как большая часть оборудования релейной защиты подстанции представляет собой терминалы MICOM серии Pх4х, целесообразно применить систему Pacis фирмы Schneider Electric в качестве SCADA-системы.

Для повышения надежности применяется топология двойное волоконно-оптическое кольцо, в которой одно кольцо является основным, а второе дублирующим. Для подключения к каждому из колец применяется два коммутатора на присоединение, каждый из которых имеет питание от независимых источников. К каждому из двух коммутаторов подключается устройство, имеющее 2 взаимозаменяемых Ethernet порта по протоколу стандарта МЭК 61850 (рисунок 3).

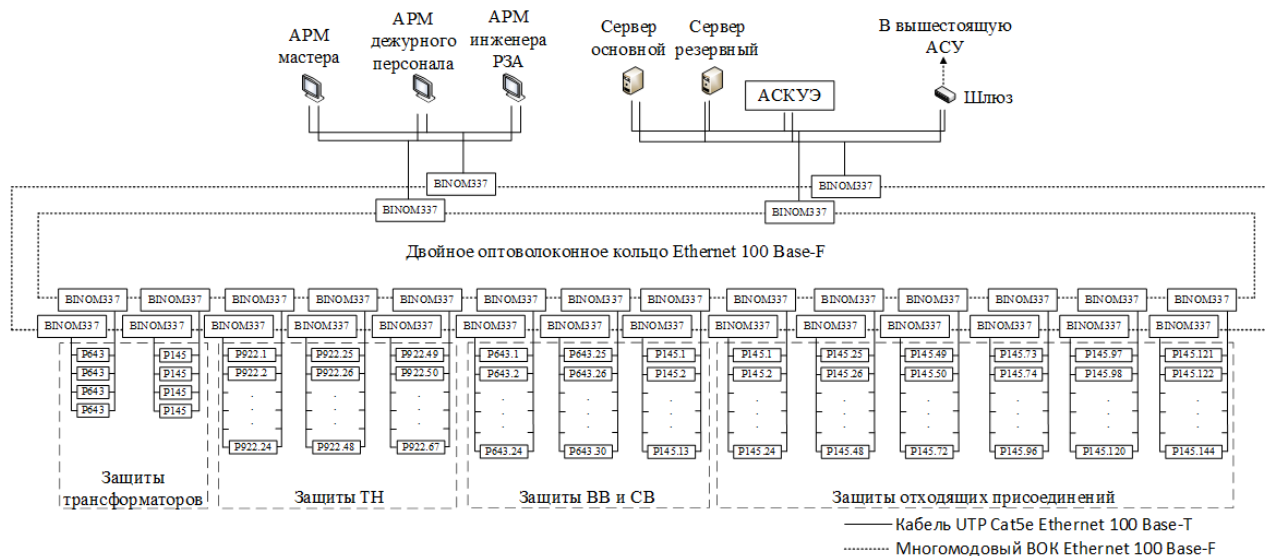


Рис. 3. Схема АСУ ТП ГПП-10

В данной работе представлен анализ действующей системы электроснабжения ГПП-10, обоснование перехода на более современные микропроцессорные средства, подбор нового оборудования дифференциальной защиты, максимальной токовой защиты, устройств автоматического повторного включения, устройств автоматического включения резерва, проанализирована электромагнитная совместимость системы, составлена структурная схема АСУ ТП ПС, АСКУЭ и составлена схема АСУ ТП ГПП-10.

Библиографический список

1. Релейная защита по применению Schneider Electric. URL: <http://www.schneider-electric.com/products/ru/ru/4600-relejnaa-zasita-po-primeneniю/>. 31.01.2021.
2. Правила устройства электроустановок. 7 изд. URL: <https://www.elec.ru/files/2016/11/02/PUE-novredaktsijaelek.pdf>. 31.01.2021.
3. Скляр В.В. Обеспечение безопасности АСУТП в соответствии с современными стандартами: методическое пособие. / В.В. Скляр. - М.: Инфра-Инженерия, 2018. – 384 с.
4. Счетчики-измерители показателей качества электрической энергии многофункциональные BINOM337. URL: <http://binom3.ru/binom337.php>. 31.01.2021.
5. Стандарт МЭК 61850. Протокол GOOSE. URL: <http://iv-sp.blogspot.com/2015/12/blog-post.html>. 31.01.2021.
6. Стандарт МЭК 61850. Протокол MMS. URL: <http://www.news.elteh.ru/arh/2013/79/04.php>. 31.01.2021.

V.N. Meshcheryakov, P.S. Ponomarev

DIGITALIZATION OF ELECTRICAL MAIN STEP-DOWN SUBSTATION-10 AT PUBLIC JOINT-STOCK COMPANY NLMK

Lipetsk state technical university

«Digital substation» allows to get one single digital data flow, which can describe status of managed object. Abstracting from the existing paradigm of building substation protection and control systems and transition to a software platform hosted on universal hardware devices and having a free distribution of functions give an opportunity to get solutions having either a fully distributed or centralized architecture. Usage a single software platform will allow to change the market of hardware devices for building substation protection and control systems and go to the market of functional algorithms.

Keywords: digital substation, relay protection and automation, automated process control system of substation, SCADA, electrosupply.

УДК 004.519

Н.Д. Рудниченко, В.В. Вычужанин, Д.С. Шибяев, Т.В. Отрадская, И.М. Петров

**РАЗРАБОТКА ПРОГРАММНОГО ОБЕСПЕЧЕНИЯ ПОДДЕРЖКИ ПРОЦЕССОВ
ПАРАЛЛЕЛЬНОЙ ОБРАБОТКИ БОЛЬШИХ ОБЪЕМОВ ДАННЫХ
О ТЕХНИЧЕСКИХ СИСТЕМАХ**

*Одесский национальный политехнический университет,
Национальный университет «Одесская морская академия»*

Задачи обработки больших объемов данных в режиме реального времени приобретают все большую актуальность в связи с повышением требований к аналитическим, управляющим и навигационным системам технических средств. В связи с этим необходимой задачей является разработка и обеспечение эффективного внедрения программных средств, позволяющих ускорить выполнение процессов обработки данных без модификации аппаратного обеспечения используемых вычислительных узлов. Возможным решением обозначенных задач является разработка программных средств и технологий поддержки процессов параллельной обработки больших объемов данных в распределенных компьютерных сетях, обеспечивающих достаточный уровень быстродействия выполняемых операций. В статье рассмотрены ключевые положения и возможности использования моделей распределенных вычислений Map Reduce, описана специфика, модульный состав и логика функционирования разработанного программного обеспечения, приведены результаты исследования его работы на тестовых выборках данных.

Ключевые слова: большие данные, параллельные вычисления, анализ данных, технические системы, Map Reduce, Hadoop

Введение

В настоящее время в информационно-технических системах, находящихся в основе аналитических инфраструктур множества современных компаний и организаций, осуществляющих свою деятельность в различных прикладных сферах бизнеса и науки, наблюдается устойчивый рост данных [1-3].

Это вызывает необходимость в разработке, внедрении и обеспечении эффективной работы информационно-вычислительных комплексов, центров обработки больших данных (ВД) и других средств и систем поддержки процессов накопления и структуризации информации [4-6].

Характер собираемых ВД является различным, в связи с чем актуальными задачами становятся определение, оценка и анализ значимых целевых критериев, с помощью которых возможно их охарактеризовать, например, по степени полезности, адекватности, целостности, ценности для бизнеса [7-9].

При этом важным фактором является быстродействие процессов обработки данных, что на практике, чаще всего, достигается путем горизонтального или вертикального аппаратного масштабирования производительности вычислительных ресурсов (серверов, распределенных узлов, вычислительных кластеров и т.д.), а также благодаря интеграции эффективных алгоритмов, программных и логических методов и моделей распределенной обработки ВД. К числу последних часто относятся решения от компаний Apache и Hadoop [10].

Следует отметить, что существующие программные решения и системы по распределенной обработке ВД от известных и крупных производителей не всегда являются экономически доступными и достаточно эффективными, не в полном объеме способными учитывать специфику необходимой предметной области, привязаны в платформе вендора и не доступны для гибкой модификации (исходные коды не являются открытыми) [11-14].

В связи с этим актуальной задачей является разработка собственного программного обеспечения (ПО) поддержки процессов параллельной обработки больших объемов данных в распределенной вычислительной сети, являющегося кроссплатформенным, основанного на модульной структуре и адаптивного для технических систем различного профиля и характера работы.

Подобная разработка возможна путем использования кроссплатформенного языка программирования Java, среды разработки IntelliJ idea и ряда дополнительных средств из Java SDK.

Описание концепции программного обеспечения

Разработанное ПО имплементирует в своем составе ряд современных технических решений и подходов к параллельной обработке данных. В частности, предусмотрена поддержка MapReduce – модели для параллельной обработки ВД в высокопроизводительных кластерных вычислительных средах.

Ее ключевыми преимуществами в рамках созданного ПО является высокая масштабируемость, обусловленная возможностью подразделения всех задач на большое число подзадач, которые могут быть выполнены параллельно на нескольких вычислительных устройствах в одном или нескольких кластерах.

MapReduce позволяет эффективно управлять вычислительными и пространственными ресурсами каждого отдельного кластера [15].

В качестве основной распределенной файловой системы (ФС) по умолчанию в ПО используется HDFS, при этом реализована поддержка GFS. Обе ФС ориентированы на интенсивные вычислительные программы, в которых распространены массивные файлы данных, оптимизированы для достижения высокой устойчивой пропускной способности лучшей поддержки пакетной обработки ВД.

В GFS используется один главный сервер GFS, который может быть размещен на выделенном узле, для координации ресурсов хранения и управления метаданными ВД.

Несколько вспомогательных серверов используются в кластере для хранения данных в виде больших блоков (фрагментов), каждый из которых имеет 64-разрядный идентификатор.

Файлы хранятся на локальном диске в виде обычных файлов, и доступ к ним осуществляется с помощью идентификатора фрагмента и смещение от начала фрагмента.

Разработанное ПО включает в свой состав 6 модулей:

1. Модуль предобработки данных (МПД). Осуществляет прием данных, их сохранение на запоминающие устройства и категоризацию (опционально).
2. Модуль конфигурации настроек (МКН) ПО. Выполняет функции параметризации внешнего вида ПО, элементов интерфейса, числа выделяемых аппаратных ресурсов на стороне сервера и локальных узлов, выбора алгоритмов расчета и задания их параметров.
3. Модуль распределения данных (МРД) по моделям и отдельным узлам в сети. Осуществляет функционал координации и активации процессов использования распределенных моделей в выбранных ФС.
4. Модуль управления вычислительными операциями (МУВО). Необходим для логирования и мониторинга процессов обработки данных, консолидирует результаты по отдельным узлам в распределенную базу данных для дальнейшего анализа.

5. Модуль оценки и анализа результатов (МОАР). Имплементирует функционал и программные зависимости обеспечения работы ряда алгоритмов по оценке производительности и эффективности выполнения вычислительных операций, в том числе методы Амдала и Густавсона, а также библиотеки PTL, MPE, Paradyн, TAU.
6. Модуль визуализации результатов (МВР) проведения вычислительных процессов.

Наибольшую ценность представляет МОАР, который может быть собран в отдельную библиотеку и подключен в необходимом рабочем окружении и среде, поддерживающей работу JVM.

Разработанная библиотека предоставляет один основной класс с набором методов для реализации необходимых параллельных конструкций. При этом все детали коммуникации, такие как управление потоками или сетевое программирование, инкапсулированы исходя из соображений безопасности.

Вместо модификации задачи для соответствия некоторой модели программирования, пользователь имеет возможность реализовывать свой алгоритм оптимальным образом. В частности, пользователь может легко реализовать такое распределение данных и работы, который лучше всего подходит для решаемой им задачи в прикладной сфере.

Разработанная библиотека предоставляет для этого необходимые инструменты, такие как нумерация потоков, передача данных и синхронизация потоков. Все коммуникации являются односторонними и асинхронными, благодаря чему процесс оценки и анализа облегчается и ускоряется.

Системы, в которые может быть подключена данная библиотека могут масштабировать посредством использования ее функционала до 5 тысяч ядер.

В разработанном ПО предусмотрен один главный узел с названием Manager. Он отвечает за установление уникальных идентификаторов задач, отправки команды для начала вычислений другой задачей, создание групп и синхронизацию всех задач. Узел Manager имеет свои собственные задачи и может выполнять операции в параллельном режиме по выбранной модели.

Мастер-узел группы всех задач имеет идентификатор группы равный 0. Каждый узел имеет свой собственный, уникальный для всех выполняемых вычислений, идентификатор. Он является физическим идентификатором или идентификатором узла.

Все узлы соединены между собой и это соединение осуществляется перед началом распределенных вычислений. На этом этапе узлы обмениваются своими физическими идентификаторами с целью координации работы и удобства мониторинга их взаимодействия МУВО.

Для использования библиотеки для параллельного выполнения, пользователю в МРД сначала нужно вызвать статический метод `start ()`, передав в него информацию об используемых классах `StartPoint` и `Storage`, а также список узлов. Список узлов используется для нумерации узлов и потоков. Каждый узел обрабатывает список для нахождения элементов данным об имени хоста - номер элемента будет использован для нумерации потоков.

Специальный узел `node0` координирует другие узлы при запуске, находящийся первым в списке. Выполнив заданные задачи каждый узел подключается к `node0` и сообщает номера элементов списка имеющимся данным об имени хоста.

Когда `node0` получает информацию о каждом узле из списка производится их нумерация числами от 0, увеличивая число на единицу при встрече каждого последующего узла, эти числа являются значением параметра `physicalId` узла. `Node0` в качестве ответа посылает всем другим узлам их физическим `physicalId`.

Следующим шагом после установления соединения каждого узла с node0 является обмен информацией между узлами и осуществления соединения каждого узла с каждым. Для этого node0 выполняет широковещательную передачу информации о каждом узле.

Широковещательная передача реализована с использованием сбалансированной древовидной структуры, в которой каждый узел содержит не более двух потомков. При запуске дерево имеет только одну вершину, а именно node0 - корень дерева.

Сообщение, передаваемое с помощью широковещательной передачи, содержит следующие данные о новом узле в дереве: physicalId; родительский physicalId; threadIds; hostname.

При получении этих данных, узел пересылает их ниже по дереву, сохраняя информацию о новом узле (в случае родительских отношений новый узел добавляется в ветвь его потомков). После этого узел устанавливает соединение с новым узлом и направляет ему информацию о себе (physicalId и threadIds).

Когда все узлы завершают этап инициализации (node0 получил подтверждение от всех узлов в распределенной сети) node0 направляет команду для запуска вычислительного процесса.

Каждый узел запускает заданное ему число потоков, каждый из которых выполняет метод переданного класса StartPoint.

Исследование результатов моделирования

Для проведения вычислительных экспериментов работы ПО выбраны 2 массива (1 и 10 Терабайт) данных о параметрах автотранспортного средства (Toyota Avensis), полученных с помощью бортового компьютера период 1 года и 10 лет. Результаты оценки скорости обработки данных приведены на рис. 1.

В обоих случаях инициализация обработки уменьшается по мере увеличения количества используемых узлов. Следует отметить, что скорость обработки BD не сильно зависит от разницы в объеме (которые отличаются на порядок) до того момента, пока число узлов составляет более 256, замедление обработки 10 ТБ составляет около 3-4 минут по сравнению с 1 ТБ.

Более значительное падение скорости вычислений и обработки данных наблюдается при уменьшении числа вычислительных узлов менее 256 и 64 (для 10 ТБ и 1 ТБ соответственно).

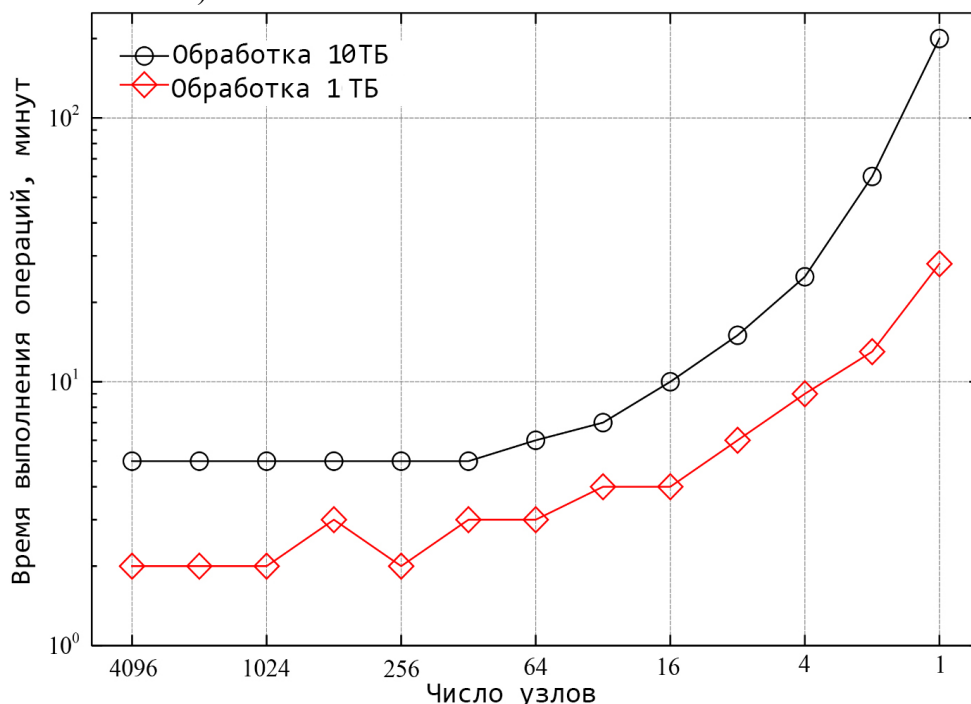


Рис. 1. Время инициализации ПО в зависимости от количества узлов

Поскольку разработанное ПО работает не в пределах одной JVM, связь между различными потоками должен быть реализована различными способами. Если потоки находятся в той же JVM, механизмы конкурентности Java могут использоваться для синхронизации и обмена информацией. Если обмен данными должен быть осуществлен между различными JVM, то должен быть использован тот или иной способ сетевой коммуникации, например, сокеты. Это необходимо для обеспечения прозрачности и эффективности обработки данных.

Заключение

Разработанное программное обеспечение показало свою работоспособность и эффективность при использовании для поддержки процессов обработки больших данных. Разработанный модуль оценки и анализа результатов может быть использован в качестве отдельной библиотеки для реализации параллельных алгоритмов, измерения из эффективности при решении вычислительных задач, а также для оценки и анализа производительности. Такая программная библиотека может быть использована в качестве функциональной зависимости для обеспечения поддержки работы распределенной системы в многопоточной вычислительной среде по различным моделям параллельного программирования, в частности, для PGAS.

Библиографический список

1. Шibaев Д.С. Оптимизация отбора и анализа информации в разноструктурных хранилищах данных / Д.С. Шibaев, В.В. Вычужанин, Н.О. Шibaева, Н.Д. Рудниченко // Информатика и математические методы в моделировании. – 2017. – №3. – С.318-324.
2. Рудниченко Н.Д. Разработка модели решающего дерева для когнитивного представления метаданных о больших объемах информации / Н.Д. Рудниченко, В.В. Вычужанин, Д.С. Шibaев, Н.О. Шibaева // Сборник материалов XXVI Международной научно-технической конференции «Информационные системы и технологии - 2020», Н. Новгород: Нижегород. гос. техн. ун-т им. Р.Е. Алексева, 2020. – С.684-670.
3. Чехарин, Е. Е. Большие данные: большие проблемы / Е. Е. Чехарин // Перспективы науки и образования. – №3 (21). – 2016. – С. 7-11.
4. Кондратьев А.А. Скрытые проблемы распределенных вычислений / А.А. Кондратьев // Электротехнические и информационные комплексы и системы. – №3. – 2015. – С. 61-65.
5. Натан М. Большие данные. Принципы и практика построения масштабируемых систем обработки данных в реальном времени / М. Натан. – М.: Вильямс, 2015. – 368 с.
6. Булыгин Н.М. Программно-аппаратный модуль параллельных вычислений / Н.М. Булыгин // StudNet. – №9. – 2020. –С. 134-143.
7. Зайнидинов Х.Н. Методы распараллеливания процессов вычисления больших объемов данных с использованием технологий параллельного программирования / Х.Н. Зайнидинов, О.У. Маллаев, Р.М. Зулунов, Ж. Нурмуродов // Автоматика и программная инженерия. – №4 (30). – 2019. – С. 81-87.
8. Биллиг В.А. Параллельные вычисления и многопоточное программирование: 2-е изд. Исправленное / В.А. Биллиг. – М.: НОУ «ИНТУИТ», 2016. –311 с.
9. Rudnichenko N. Decision Support System for the Machine Learning Methods Selection in Big Data Mining / N. Rudnichenko, V. Vychuzhanin, I. Petrov, D. Shibaev // Proceedings Of The Third International Workshop on Computer Modeling and Intelligent Systems (CMIS-2020): session 6 “Intelligent Information Technologies” April 27-May 1, 2020. – Zaporizhzhia: NU “Zaporizhzhia Polytechnic” (edited by S. Subbotin), 2020. – P. 872-885.

10. Shibaev D. S. Data control in the diagnostics and forecasting the state of complex technical systems / D. S. Shibaev, V. V. Vyuzhuzhanin, N. D. Rudnichenko, N. O. Shibaeva, T. V. Otradsкая // Herald of Advanced Information Technology. – 2019. – Vol. 2. – № 3. – PP. 183-196
11. Вычужанин В.В. Методы информационных технологий в диагностике состояния сложных технических систем. Монография / В.В. Вычужанин, Н.Д. Рудниченко. – Одесса: Экология, 2019. – 178 с.
12. Рудниченко Н.Д. Разработка концепции модуля интеллектуального анализа больших объемов данных в транспортных системах на базе методов машинного обучения / Н.Д. Рудниченко, В.В. Вычужанин, Н.О. Шибалева, Д.С. Шибалев, Н.И. Гежа // Межвузовский сборник научных статей (с международным участием) «Актуальные проблемы автотранспортного комплекса» [отв. ред. О.М. Батищева]. – Самара: Самар. гос. техн. ун-т, 2020. – С. 85-92.
13. Vychuzhanin V. V. Analysis and structuring diagnostic large volume data of technical condition of complex equipment in transport / V.V. Vychuzhanin, N.R. Rudnichenko, Z. Sagova, M. Smieszek, V.V. Cherniavskyi, A.I. Golovan, M.V. Volodarets // 24th Slovak-Polish International Scientific Conference on Machine Modelling and Simulations - MMS 2019, 3-6 September 2019, Liptovský Ján, Slovakia.
14. Rudnichenko N. Complex Technical System Condition Diagnostics and Prediction Computerization / N. Rudnichenko, V. Vychuzhanin, A. Polyvianchuk, V. Mateichyk // Proceedings Of The Third International Workshop on Computer Modeling and Intelligent Systems (CMIS-2020): session 6 “Intelligent Information Technologies” April 27-May 1, 2020. – Zaporizhzhia: NU “Zaporizhzhia Polytechnic” (edited by S. Subbotin), 2020. – P. 42-56.
15. Гладкий М.В. Модель распределенных вычислений MapReduce / М.В. Гладкий // Труды БГТУ. Серия 3: Физико-математические науки и информатика. – №6 (188). – 2016. – С. 194-198.

N. D. Rudnichenko, V.V. Vychuzhanin, D.S. Shibaev, T.V. Otradsкая, I.M. Petrov

DEVELOPMENT OF SOFTWARE TO SUPPORT PARALLEL PROCESSING OF BIG DATA ON TECHNICAL SYSTEMS

*Odessa National Polytechnic University,
National University "Odessa Maritime Academy"*

The tasks of processing large amounts of data in real time are becoming increasingly important due to the increasing requirements for analytical, control and navigation systems of technical means. In this regard, a necessary task is to develop and ensure the effective implementation of software tools that can speed up the execution of data processing processes without modifying the hardware of the computing nodes used. A possible solution to these problems is the development of software tools and technologies for supporting the processes of parallel processing of large amounts of data in distributed computer networks, which provide a sufficient level of performance of the operations performed. The article discusses the key provisions and possibilities of using distributed computing models Map Reduce, describes the specificity, modular composition and logic of the functioning of the developed software, presents the results of the study of its work on test data samples.

Keywords: big data, parallel computing, data analysis, technical systems, Mapreduce, Hadoop.

РОБОТИЗИРОВАННАЯ СИСТЕМА ТРАНСПОРТИРОВКИ ИЗДЕЛИЙ ПРОИЗВОДСТВА

¹*Белорусский национальный технический университет, г. Минск*

²*Белорусский государственный университет информатики и радиоэлектроники, г. Минск*

Рассматривается микропроцессорное управление многозвенным роботом-манипулятором, обеспечивающим транспортировку изделий производства в соответствии с заданным технологическим процессом. Излагается методика разработки прикладного программного обеспечения функционирования многоприводной мехатронной системы под управлением промышленного микроконтроллера фирмы OMRON. Электромеханическая конструкция манипулятора позволяет обеспечить высокоточное позиционирование рабочего органа. Представлен внешний вид разработанного робота - манипулятора и его применение в технологическом процессе быстрой термической обработки кремниевых пластин в электронной отрасли промышленности. Разработанная мехатронная система робота- манипулятора может быть технологически ориентирована на решение других задач промышленного производства.

Ключевые слова: мехатронная система, робот-манипулятор, прикладная программа, технологический процесс, программируемый логический контроллер, программируемый терминал, микроконтроллер, электропривод.

Одним из главных направлений, определяющих научно-технический уровень развития современного производства, является комплексная автоматизация технологических процессов с применением высокотехнологичного и производительного оборудования с требуемым уровнем загрузки, скоростной транспортировки и синхронизации подачи изготавливаемых изделий к рабочим позициям во времени последовательно, параллельно или параллельно-последовательно в соответствии с заданным технологическим циклом. На разных стадиях этого цикла требуется использование роботизированных систем транспортировки производственных объектов. Основным критерием таких систем является точность позиционирования и соблюдение ограничений на перемещение, так как операция может выполняться в ограниченной области пространства.

Целью данной статьи является автоматизация проектирования мехатронной системы робота-манипулятора типа Scara под управлением промышленного программируемого логического контроллера OMRON для транспортировки изделий производства.

Для реализации поставленных задач была выбрана конструкция манипулятора, представленная на рис.1.

Данная конструкция разработана республиканским унитарным предприятием «Политехник» и представляет собой двухзвенный механизм с разомкнутой кинематической цепью, оснащенный тремя сервоприводами и рабочим органом – схватом (на рисунке не показан). С помощью сервопривода, расположенного в основании, осуществляется вертикальное перемещение рабочего органа. Два других сервопривода изменяют угловые положения звеньев и, тем самым, обеспечивают позиционирование рабочего органа в горизонтальной плоскости. Поворот кисти механически синхронизирован с поворотом предплечья. Для захвата и выгрузки объектов предусмотрен переворот схвата на 180 градусов.

Приведенная конструкция манипуляционного механизма позволяет автоматизировать доставку заготовок, деталей, других объектов к обрабатывающим узлам

(постам), расположенным в его рабочей зоне. Рабочая зона – это пространственная область, ограниченная рабочим ходом в вертикальном и длинами звеньев манипулятора в горизонтальном направлениях. Кроме этого, на рабочую зону также влияют кабелеукладчики.

Работа робота-манипулятора осуществляется в соответствии с траекторией движения, определяемой технологическим процессом, и задается множеством пронумерованных точек (номеров производственных обрабатывающих постов), например, как показано на рис.2. На этом рисунке приведены два основных положения звеньев манипулятора, обеспечивающих изменение положения рабочего органа - одно для совершения поворота, второе - для линейного позиционирования в заданную точку. Производственные обрабатывающие посты отмечены на рисунке жирными кружочками большего радиуса.

Алгоритм работы робота-манипулятора приведен в [1].



Рис. 1. Конструкция манипуляционного механизма

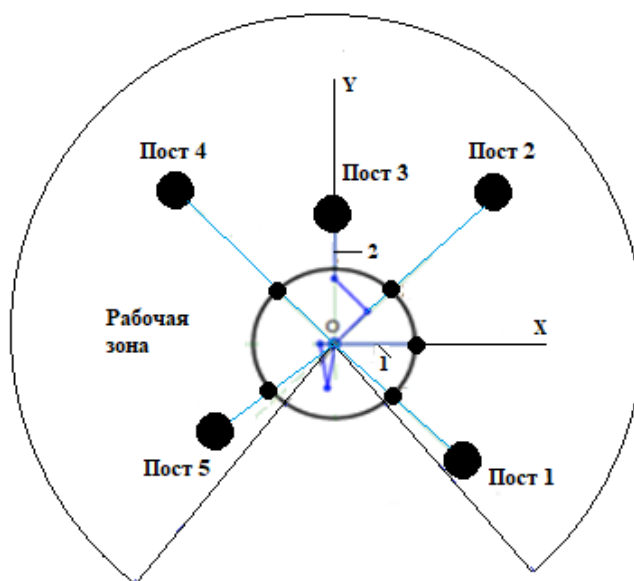


Рис. 2. Множество точек для задания траектории движения рабочего органа робота (1-положение робота при повороте, 2-положение робота при линейном движении вдоль радиуса)

Для автоматизации управления перемещением объектов производства по описанному выше технологическому маршруту использован программируемый

логический контроллер (микроконтроллер) фирмы OMRON, широко применяемый в настоящее время при решении различных задач в промышленности.

Информационная система организации взаимодействия микроконтроллера с роботом-манипулятором, с интегральной средой создания управляющей программы и с программируемым терминалом, используемым для ручного управления манипулятором [2], приведена на рис.3.

На этом рисунке мехатронная система робота-манипулятора представлена двумя блоками (на рисунке справа) – непосредственно микроконтроллером и электромеханической составляющей самого манипулятора. Функционирование мехатронной системы по выполнению последовательных действий, определяемых технологическим процессом, предполагает создание прикладной программы, состоящей из двух уровней – верхнего и нижнего.

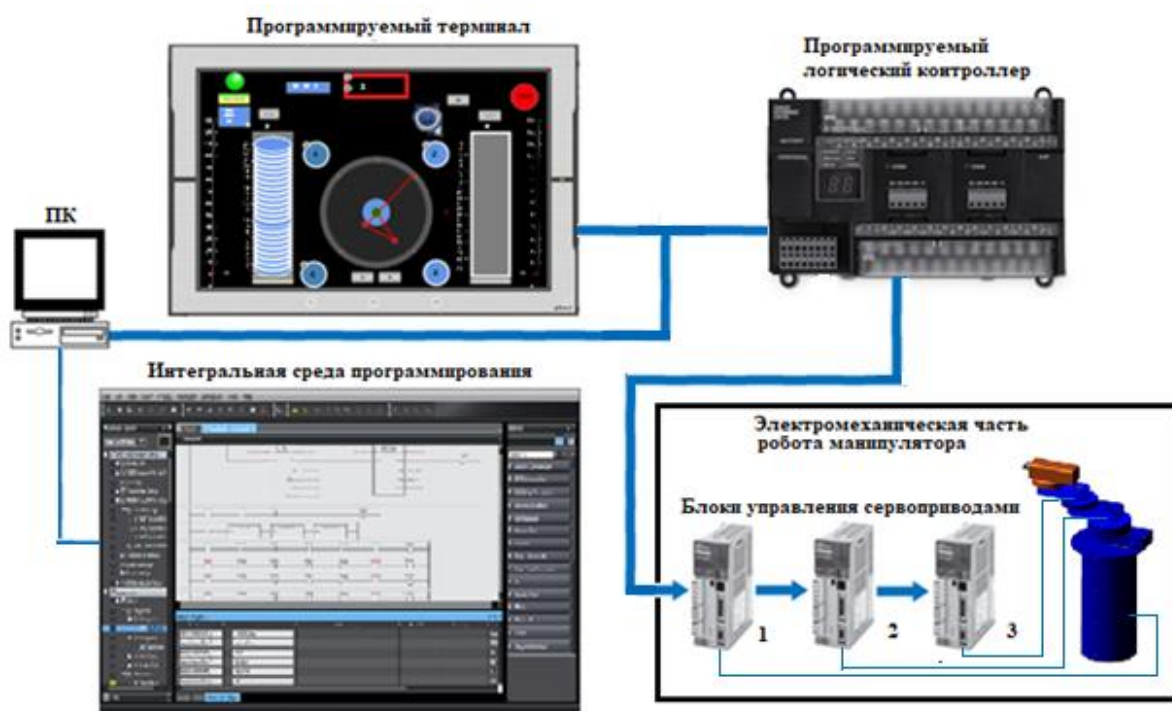


Рис. 3. Информационные связи мехатронной системы робота-манипулятора и среда разработки программного обеспечения

Верхний – технологический, являющийся внешним уровнем относительно электропривода, и нижний – внутренний, непосредственно управляет электроприводом. На верхнем уровне программа должна обеспечить формирование технологического задания по управлению рабочим органом робота-манипулятора в соответствии с заданным технологическим маршрутом (рис.2).

Задачей нижнего уровня является формирование статических, динамических характеристик непосредственно электропривода (двигатель, механическая передача, рабочий орган). В этом случае рабочий орган, физически принадлежащий манипулятору, рассматривается как составная часть электропривода, т.е., программа второго уровня содержит свои связи, в том числе, и обратные, и описывает поведение электропривода под воздействием входных сигналов от системы управления первого уровня. Оценить качество электропривода невозможно без совместного статического и динамического взаимодействия составных частей данной системы. В такой постановке, электропривод является объектом исследования и проектирования.

Для программирования микроконтроллеров OMRON с учетом двухуровневой структуры управляющей прикладной программы используется интегральная среда Sysmac

Automation Studio, поддерживающая язык релейно-контактной логики, алгоритмический язык в виде структурированного текста, функций и функциональных блоков. Кроме этого, она имеет развитые возможности редактирования и отладки программ, программирование переносного терминала для ручного управления технологическим процессом, средства формирования структурной схемы на базе различных устройств (двигателей, датчиков и т.д.), поддержку программирования с использованием эмуляторов, расширенную библиотеку стандартных функций и функциональных блоков, трансляцию исходного кода в исполняемый и «прошивку» его в микроконтроллер.

Используя описанные выше возможности интегральной среды программирования, была разработана прикладная программа микроконтроллерного управления роботом-манипулятором, структурная схема которой приведена на рис.4.



Рис. 4. Структурная схема управляющей программы

Управляющая программа обеспечивает связь отдельных модулей и координирует их работу. *Модуль управления роботом* организует перемещения рабочего органа робота вдоль осей X, Y, Z. *Модуль задания координат* позиционных точек представлен в виде пользовательской функции формирования массива, содержащего декартовы координаты характерных точек, описывающих позиции расположения постом, обслуживаемых роботом. *Модуль описания команд* содержит набор команд для позиционирования рабочего органа робота в требуемые точки декартового пространства, заданные в модуле координат. *Модуль управления приводами* содержит стандартные блоки библиотеки OMRON и непосредственно связан с сервоприводами, изменяющими физическое положение звеньев робота. *База данных* представляет собой список переменных, массивов и структур для хранения геометрических параметров, параметров движения, свойств команд и другой информации. *Функции внешнего управления* позволяют установить автоматический или пошаговый режим выполнения команд, действий по контролю точности позиционирования и др.

Следует отметить, что несмотря на достаточно высокие возможности интегрированной среды программирования, разработка прикладной программы обслуживания технологического оборудования роботизированными устройствами занимает значительное время в общем процессе создания автоматизированной системы. Это обусловлено необходимостью постоянной связи с реальным физическим контроллером и роботом-манипулятором для непосредственного тестирования отлаживаемого программного обеспечения в реальном режиме времени. Как правило, на начальных стадиях создания автоматизированной системы есть только проект, а оборудование появляется в процессе работы.

Чтобы сократить время разработки программной поддержки микропроцессорного управления роботом-манипулятором, в настоящей работе была реализована упрощенная

2D модель манипулятора с визуализацией текущего геометрического расположения его звеньев при выполнении технологических операций по транспортировке объектов производства в соответствии с рис.2. Отображение модели производится на экране программируемого терминала (рис.3) и совмещено с интерфейсной частью управления работой робота в общем технологическом цикле. В дальнейшем, на стадии эксплуатации модель копирует движения реального робота и служит источником информации для управления и его состояния.

В этом случае, процесс отладки программы для микроконтроллера не связан с физическим наличием микроконтроллера и терминала, так как интегрированная среда программирования позволяет запустить обе программы (для микроконтроллера и терминала) синхронизировано в режиме эмуляции.

Эмулятор контроллера достаточно полно воспроизводит выполнение основных команд программы, связанных с обработкой временных состояний работы электрических приводов, таймеров, линий задержек и других используемых элементов. В связке с визуализацией поведения виртуального робота-манипулятора на экране эмулятора терминала эффективность отладки и тестирования программного обеспечения значительно повышается. Кроме этого, исключается возможность механического повреждения манипулятора и опасность травмирования людей при тестировании программного обеспечения и отработке алгоритмов управления и позиционирования рабочего органа.

Для программирования анимационной модели манипулятора на панели терминала интегрированная среда предоставляет минимальный набор простейших геометрических фигур и язык программирования Visual Basic. С помощью этих средств имеется возможность реализовать простейшую графическую поддержку модели робота-манипулятора. Для создания 3D модели [3] требуются дополнительные библиотеки.

Изложенное выше было реализовано при разработке программного обеспечения автоматизации транспортных операций мехатронной системой на базе двухзвенного робота-манипулятора при организации технологического процесса быстрой термической обработки кремниевых пластин [4] от участка загрузки (подающей кассеты) к пункту геометрической ориентации, далее к реактору, в котором производится непосредственно операция отжига. После этого робот доставляет пластину к месту охлаждения и далее в конечную точку технологического процесса (приемную кассету). Кроме указанных операций, программа микроконтроллера выполняет действия, связанные с картированием подающей и приемной кассет с целью их наличия/отсутствия или правильного расположения с помощью встроенного лазерного датчика, а также базированием, обеспечивающих контроль текущих параметров, заданным при настройке робота [5].

Управление роботом-манипулятором (рис.3), расположение постов обработки пластин, положение звеньев и рабочего органа в каждый момент времени осуществляются графической моделью и интерфейсными элементами терминала в режиме реального времени.

Программное обеспечение поддерживает два режима организации технологического процесса: последовательный (рассмотрен выше) и произвольный (параллельный). Режим параллельной обработки требует установки на постах датчиков, сигнализирующих об окончании выполненных операций на соответствующих постах и наличии других дополнительных условий о возможности принятия пластины следующим постом.

Внешний вид установки, обслуживаемой мехатронной системой с манипуляционным механизмом, показан на рис.5.

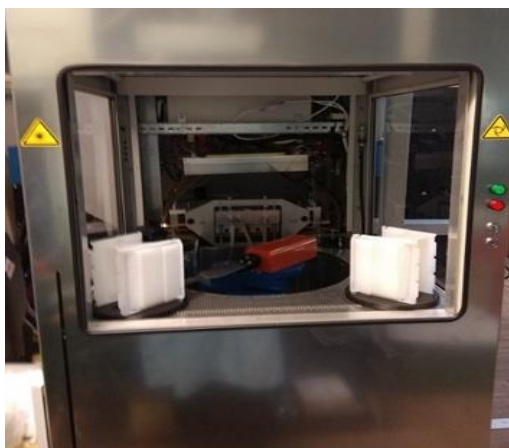


Рис. 5. Роботизированная установка быстрой термической обработки кремниевых пластин

Выводы

1. Разработано программное обеспечение высокоточного позиционирования робота-манипулятора на базе двухзвенного механизма под управлением промышленного программируемого логического контроллера OMRON.
2. Предложена методика разработки программы микроконтроллерного управления роботизированным устройством, позволяющая сократить временные затраты при создании автоматизированных систем.
3. Разработана мехатронная система робота-манипулятора для обслуживания процесса быстрой термической обработки кремниевых пластин, которая может быть технологически ориентирована на решение других задач промышленного производства.

Библиографический список

1. Гурский Н.Н. Математические и компьютерные модели мехатронных систем аддитивного производства // Математические методы в технике и технологиях. Сборник трудов международной научной конференции. Том 12. Часть 1. – Санкт-Петербург, 2017. – С. 151-156.
2. Зенкевич С. Л. Основы управления манипуляционными роботами / С. Л. Зенкевич, А. С. Ющенко // М.: Изд-во МГТУ им. Н. Э. Баумана, 2004. – 480 с.
2. Гурский Н.Н. Управление мехатронной системой на базе многосвязных роботов манипуляторов / Н.Н. Гурский, Ю.А. Скудняков, В.С. Артющик, А.Н. Безручко // Наука и техника. Международный научно-технический журнал. Том 18, № 4. – Минск, 2019. – С. 350 - 354.
3. Солодуха В.А. Роботизированная установка быстрой термической обработки для создания изделий электронной техники / В.А. Солодуха, В.А. Пилипенко, В.П. Яковлев // Доклады БГУИР. №4 (122). – Минск, 2019. – С. 92-97.
4. Гурский Н.Н. Автоматизация транспортировки объектов производства роботоманипулятором и диагностика его настроек / Н.Н. Гурский, В.С. Артющик // Математические методы в технике и технологиях. Сборник трудов международной научной конференции. Том 12. Часть 3. – Санкт-Петербург, 2020. – С. 16 -19.

N.N. Hurski¹, Y.A. Skudnyakov²

ROBOTIC TRANSPORTATION SYSTEM PRODUCTS OF PRODUCTION

¹Belarusian national technical University

²Belarusian State University of Informatics and Radioelectronics

Microprocessor-based control of a multi-link robot by a manipulator is considered, which ensures the transportation of production products in accordance with a given technological process. The article describes a method of developing applied software for the functioning of a multidrive mechatronic system under the control of an industrial microcontroller from OMRON. The electromechanical design of the manipulator allows for high-precision positioning of the working body. The appearance of the developed robot manipulator and its application in the technological process of fast heat treatment of silicon wafers in the electronics industry are presented. The developed mechatronic system of the robotic arm can be technologically oriented towards solving other problems of industrial production.

Keywords: robotic system, 3d objects, the positioning mechanisms, differential equations, electric drive, an additive process.

ПРИМЕНЕНИЕ СРЕДСТВ АВТОМАТИЗАЦИИ ОБРАЗОВАТЕЛЬНОГО ПРОЦЕССА ДЛЯ УЛУЧШЕНИЯ КАЧЕСТВА ПОДГОТОВКИ СПЕЦИАЛИСТОВ

Белорусский государственный университет информатики и радиоэлектроники, г. Минск

В данной работе предложен подход по организации образовательного процесса с применением средств автоматизации для улучшения качества составляемого и усваиваемого материала.

Ключевые слова: образование, информационные технологии, социологический опрос, алгоритмы, большие данные.

Введение

Задачами данного исследования являлись: исследование проблем образования, формирование гипотез и разработка решений, соответствующих современным требованиям. Образование является ключевой сферой, формирующей государственное устройство, благодаря которому закладываются экономические, политические и социальные институты общества. Конкурентоспособность страны зависит от уровня подготовки специалистов. Именно они определяют возможности государства выполнять свои обязательства перед своими гражданами как внутри страны, так и на международной арене. На сегодняшний день перед странами СНГ стоит серьезная задача: подготовить поколение квалифицированных кадров, создать достойную систему мотивации и возможности для самореализации. Эти шаги позволят минимизировать возможность оттока специалистов достаточно высокого профессионального уровня за рубеж. Если в конце 20-го века главным ресурсом экономики считалось сырье, то уже в начале 21-го приоритет начал меняться в пользу информационных ресурсов. Подтверждением вышеупомянутого тезиса является следующая цифра: за 4-й финансовый квартал 2015-года компания Apple достигла рекордной выручки в 234 миллиарда долларов США.

Использование образовательного стандарта высшего образования первой ступени Республики Беларусь потенциально позволяет обеспечивать формирование следующих групп компетенций:

- академических компетенций, включающих знания и умения по изученным
- дисциплинам, способности и умения учиться;
- социально-личностных компетенций, включающих культурно-ценностные
- ориентации, знание идеологических, нравственных ценностей общества и государства и умение следовать им;
- профессиональных компетенций, включающих способность решать задачи,
- разрабатывать планы и обеспечивать их выполнение в избранной сфере профессиональной деятельности.

Формирование и развитие социально-профессиональной, практико-ориентированной компетентностей позволяют сочетать базовые профессиональные и специализированные знания для решения задач в сфере профессиональной и социальной деятельности.

Таким образом, основной целью профессионального образования является подготовка квалифицированного работника, конкурентоспособного на рынке труда, компетентного, ответственного, свободно владеющего своей профессией и ориентирующегося в смежных областях деятельности. Он должен быть способен к эффективной работе по специальности на уровне мировых стандартов и готов к постоянному профессиональному росту. Немаловажное значение имеет его способность к социальной и профессиональной мобильности.

Для более подробного изучения ситуации, связанной с достижением приведенных выше целей, было решено провести опрос как среди нынешних студентов, так и людей, которые уже получили высшее или среднее специальное образование. Решение данной задачи осуществлялось с помощью разработанного в работе программно-алгоритмического обеспечения обработки больших массивов данных.

Оценка эффективности и алгоритмизация процесса обучения

Отличительными чертами качественной современной системы образования являются: качество предоставляемой информации и эффективность методики её подачи. В результате влияния этих факторов студент будет иметь максимальную заинтересованность к образовательному процессу.

Данные о наличии в образовании неактуальных сведений представлены на рис. 1. Как видно из рисунка по результатам опроса, подавляющее большинство (78,6%) респондентов в ходе учебы испытывали ощущение неактуальности преподаваемых знаний. Столь высокий показатель неудовлетворенности получаемых знаний может быть связан с тем, что часть учащихся во время обучения уже имеют рабочее место, где они независимо от ВУЗа получают самые современные профессиональные навыки, которые дают возможность заработка и осваивают наиболее актуальные методы работы. Следует также учитывать, что трудоустроенный в современной компании человек может оценивать рыночную перспективу получаемых в ВУЗе сведений. В этом случае системе образования, чтобы успешно конкурировать с компанией-работодателем и обеспечивать самой актуальной информацией, необходимо прилагать большие усилия по разработке и применению в учебном процессе электронных средств обучения, содержащих все необходимые виды обеспечения для подготовки конкурентных специалистов. Следует учесть, что чрезмерный меркантилизм может приводить к результату, когда студенты ошибочно полагают, что совершенствовать необходимо только те знания, которые приносят им деньги. Практика современной наукоемкой экономики показывает, что достижение наибольшего успеха можно получить в том случае, когда используются системные знания и эффективные результаты, полученные в разных сферах человеческой деятельности. При использовании такого подхода возможен наибольший научно-технический и коммерческий успех.

Решением вышеперечисленных проблем может быть использование в образовательной программе фундаментальных знаний как по профилю подготовки, так и по иным смежным областям. Их понимание является весьма актуальным.

Испытавали ли вы ощущение неактуальности знаний, преподаваемых в вашем учебном заведении?

14 ответов

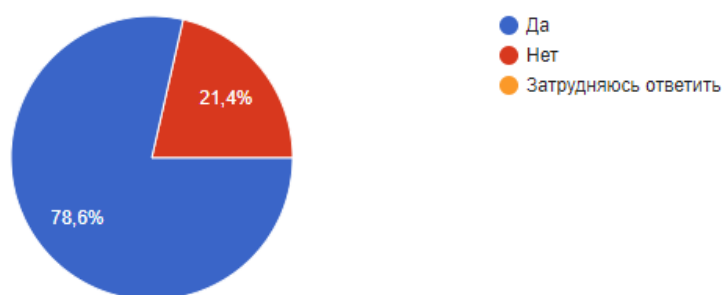


Рис. 1. Результат опроса о неактуальности преподаваемых знаний

Неактуальность преподаваемых знаний напрямую влияет на заинтересованность студента в обучении, что многократно снижает эффективность усваиваемой информации и нивелирует старания преподавательского состава к формированию учебного материала и улучшению качества его преподавания. Согласно рис. 2 частое отсутствие заинтересованности к обучению испытывало 21,4% респондентов.

Испытывали ли вы периоды отсутствия заинтересованности?

14 ответов



Рис. 2. Результат опроса о периодах отсутствия заинтересованности обучаемых

Ещё 71,4% опрошенных сталкивались с подобными периодами реже. При интерпретации данных результатов необходимо учитывать разный уровень подготовки и скорости усваивания профильной информации студентами, обусловленный различиями в способности усваивать учебный материал. Для поддержания уровня заинтересованности у разных групп на одинаковом уровне в работе предлагается ввести индивидуальную рекомендательную систему, которая представлена на рис. 3.

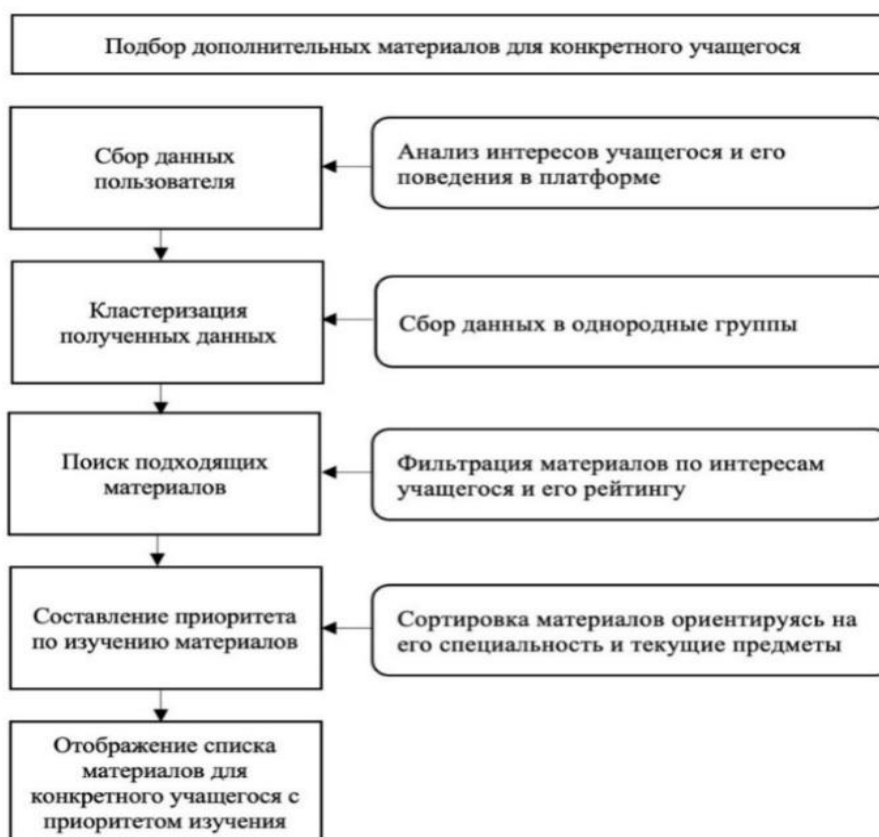


Рис. 3. Схема работы алгоритма по подбору дополнительных материалов

Также для увеличения мотивационной составляющей обучения необходимо внедрение алгоритма геймификации на основе рейтинговой системы. Система рейтинга учащегося будет базироваться на действиях пользователя на платформе: участие в тематических мероприятиях, благотворительности, учет среднего балла, активность в научно-практических конференциях. Можно предполагать, что вышеупомянутый алгоритм

будет благотворно сказываться [1] на уровне вовлеченности студентов в учебный процесс, а также предоставит возможность составления “портрета пользователя” [2]. В будущем портрет пользователя сможет сыграть ключевую роль как в улучшении образовательного процесса, так и в улучшении процента трудоустроенности. Он позволит компаниям получать более детальную информацию об активности кандидата во время образовательного процесса, уровне его вовлеченности, качестве выполненных во время обучения проектов, интересах и перспективах. Таким образом, университет получает дополнительную возможность извлечения прибыли, потому что поиск сотрудника является трудозатратным процессом для компаний.

Как было показано выше, 71,4% обучаемых испытывали периодические спады в заинтересованности предоставляемой информации, что является проблемой, требующей решения. Для поддержания вовлеченности студентов на всем протяжении образовательного процесса, на платформе разработанного в работе программно-алгоритмического обеспечения можно внедрить групповые формы обучения для объединения учащихся по интересам. Это позволит эффективнее организовывать деятельность студентов, направлять их знания и энергию на взаимовыгодные проекты, научные конференции и иную практико-ориентированную деятельность.

Одной из главных задач образовательных учреждений является обеспечение информационной чистоты, т.е. предоставление достоверной, грамотно структурированной информации. Для решения данной проблемы необходимо организовать беспрепятственный доступ к предоставляемым образовательным материалам. Данные, характеризующие наличие трудности с доступом к информации, представлены на рис. 4.

Испытывали ли вы трудности с доступом к информации, которая предоставляется лектором?

14 ответов

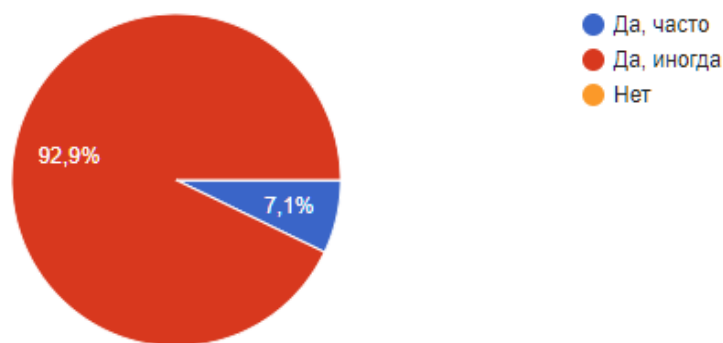


Рис. 4. Результат опроса о трудностях с доступом к информации

Согласно результатам социологического опроса, 7,1% респондентов часто испытывали проблемы с доступом к учебным материалам, 92,9% испытывали схожие проблемы, с меньшей регулярностью. Предложенная в работе платформа позволяет решить проблему с прозрачностью предоставляемых материалов путем их интерактивного предоставления до и после занятий с целью первичного и повторного ознакомления.

Наличие материалов в электронном виде также открывает возможности для интеграции в систему алгоритмов автоматического составления проверочных заданий после занятий. Таким образом, лучше происходит закрепление усвоенного материала. Кроме этого, следует учитывать, что система может предлагать вопросы для проверки не только на основании анализа материалов, предлагаемых преподавателем. Она также может использовать сформированный “портрет пользователя” и доступ в Интернет для обеспечения обучаемого индивидуальными проверочными заданиями. Это может увеличить вовлеченность, поскольку будут учтены интересы студентов.

Платформа, изучая поведение пользователя в ней, с помощью анализа большого массива данных имеет возможность подбирать материал под конкретного ученика при формировании дополнительной литературы из интернет-ресурсов [3]. При успешно составленном рекомендательном алгоритме большее количество студентов будет заинтересовано в самообразовании и изучении дополнительных материалов. При этом, нагрузка на преподавателя снизится, так как рекомендации могут быть составлены системой с помощью алгоритмов.

Для дополнительного стимулирования интереса к получению студентами знаний их деятельность на платформе будет прямо коррелировать с размером потенциальной финансовой скидки на обучение. Алгоритм её вычисления можно будет настроить более гибко, чем при использовании подхода, основанного на среднем балле и посещаемости. Теперь можно будет оценивать вклад отдельно взятого учащегося для учебного заведения, государства и общества.

Общая схема алгоритма организации образовательного процесса представлена на рис. 5.

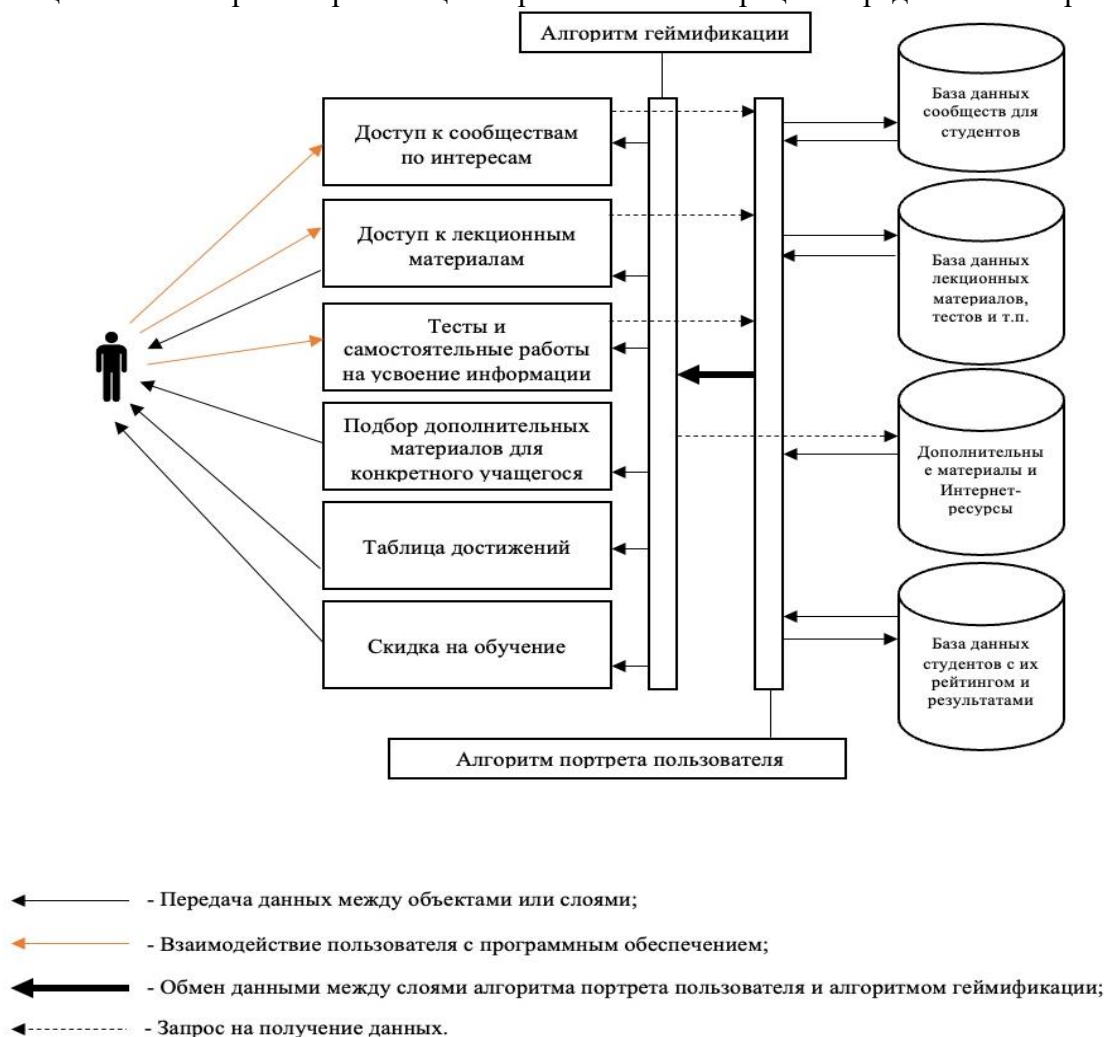


Рис. 5. Схема работы алгоритма организации образовательного процесса

Заключение

Автоматизация образовательного процесса позволяет повысить качество подготовки будущих профессионалов, что напрямую связано с улучшением макроэкономических, экономических показателей, которые способствуют формированию стабильного социума, политической обстановки и положения страны на международной арене.

Внедрение и развитие предложенного в рамках исследования подхода и средств автоматизации организации образовательного процесса позволит существенно повысить как уровень заинтересованности обучающихся, так и качество преподаваемого материала.

Применение алгоритма геймификации, интерактивного подхода организации процесса обучения посредством интеграции тематических групп, методики подбора индивидуального материала с помощью алгоритма рекомендательной системы позволяет обеспечить требуемый уровень автоматизации, гибкости и, в целом, высокую эффективность современного процесса подготовки высококвалифицированных специалистов.

Библиографический список

1. Капкаев Ю.Ш. Геймификация образовательного процесса / Ю.Ш. Капкаев,
2. Лешина В.В., Бенц Д.С. // Журнал ВАК: «Проблемы современного педагогического образования». – Минск, 2019. – С. 213 – 215.
3. Создание портрета целевой аудитории // ХабрХабр [Электронный ресурс].
4. Электронные данные. – Режим доступа: <https://habr.com/ru/company/netologyru/blog/346702/> (дата обращения 10.01.2021).
5. Рекомендательные системы // ХабрХабр [Электронный ресурс]. – Электронные данные. – Режим доступа: <https://habr.com/ru/company/lanit/blog/420499/> (дата обращения 10.01.2021).

Y.A. Skudnyakov, I.A. Karpekin, N.A. Poluectov

APPLICATION OF EDUCATIONAL PROCESS AUTOMATION TOOLS TO IMPROVE THE QUALITY OF SPECIALIST TRAINING

Belarusian state university of informatics and radioelectronics, Minsk

In this paper, we propose an approach to the organization of the educational process with the use of automation tools to improve the quality of the compiled and assimilated material.

Keywords: education, information technologies, sociological survey, algorithms, big data.

ПРИКЛАДНЫЕ ПРОГРАММНЫЕ КОМПЛЕКСЫ В СОСТАВЕ ПЛАТФОРМЫ «ЦИФРОВОЕ ПРЕДПРИЯТИЕ»

Российский федеральный ядерный центр – Всероссийский научно-исследовательский институт экспериментальной физики

В статье рассматриваются комплексы прикладного программного обеспечения, входящие в состав и функционирующие в рамках информационно-технологической платформы «Цифровое предприятие».

Ключевые слова: прикладное программное обеспечение, программный комплекс, цифровое предприятие, информационно-технологическая платформа.

Динамика производства и внешней среды все чаще и чаще приводит научно-производственные предприятия, выпускающие сложную наукоемкую продукцию, к пониманию необходимости внедрения концепции «Цифровое предприятие» - концепции комплексной поэтапной автоматизации современного предприятия.

Концепция (модель) «Цифровое предприятие» – это комплексная информационно-технологическая платформа, обеспечивающая информационную поддержку всех этапов жизненного цикла изделий научно-производственного предприятия, а также автоматизацию всего процесса создания изделия от управления требованиями до выпуска конечной продукции, ее гарантийного обслуживания и авторского надзора.

Электронная модель «Цифровое предприятие» может применяться как виртуальный (электронный) аналог реального предприятия для анализа возникающих проблем, проведения исследований и экспериментов для повышения качества продукции, сокращения сроков выполнения процессов и сокращения материальных затрат

Основа построения электронной цифровой модели предприятия – комплексная процессная модель предприятия, применяемая для наглядного и целостного представления деятельности конкретного предприятия, сформированная в ходе оптимизации как основных, так и вспомогательных, обеспечивающих бизнес-процессов предприятия, взаимоувязки моделей процессов основной и обеспечивающей деятельности с моделями организационной структуры, компетенциями, моделями целей, научно-производственной базы.

Комплексная информационно-технологическая платформа «Цифровое предприятие» с входящей в него системой промышленной автоматизации, основанная на комплексной процессной модели предприятия, формирует единое непрерывное сквозное информационное пространство ведения как научно-производственной, так и поддерживающей деятельности предприятия с едиными правилами управления жизненным циклом изделия, см. рис.1.

Однако платформа «Цифровое предприятие» это не просто «модное» направление исследования, не абстрактное понятие, а сложный многокомпонентный комплекс (система) взаимодействующих в едином информационном пространстве конкретных цифровых продуктов – прикладного программного обеспечения (прикладного ПО), баз данных, автоматизированных и информационных систем.

Помимо крупных автоматизированных информационно-программных систем предприятия (таких, как PDM и ERP системы, рассмотрение которых требует изложения в отдельной статье/докладе), в состав платформы «Цифровое предприятие» входят комплексы прикладного программного обеспечения, функционирующие в едином информационном пространстве ведения деятельности предприятия и обеспечивающие управление ресурсами предприятия, производством, безопасностью и т.п.

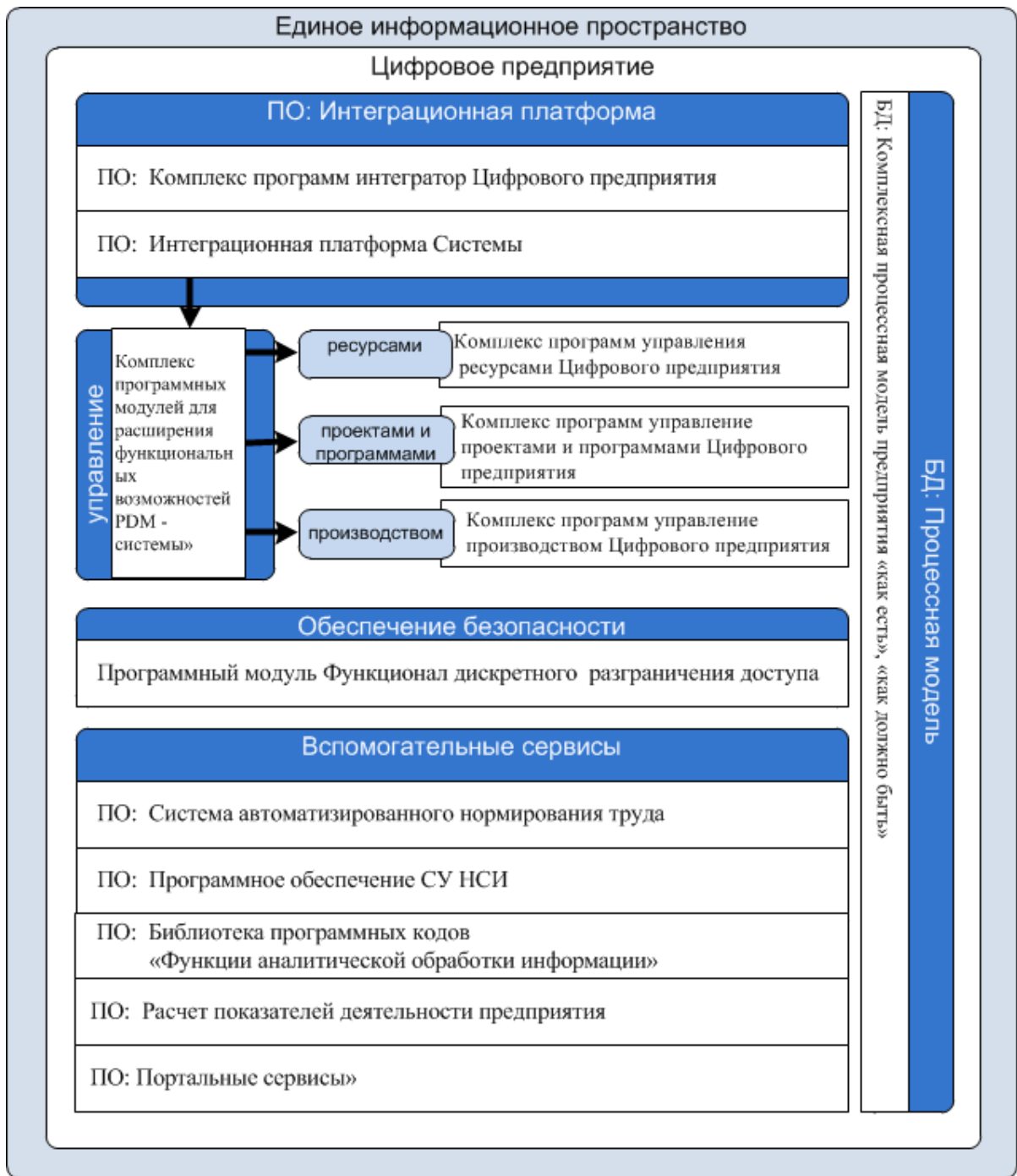


Рис. 1. Единое программное информационное пространство предприятия

Рассмотрим некоторые комплексы прикладного программного обеспечения, разработанные и функционирующие в рамках концепции «Цифровое предприятие» (местоположение рассматриваемых комплексов в составе платформа «Цифровое предприятие» приведено на рис.1).

1. Комплекс программ Интегратор Цифрового предприятия.

Предназначен для интеграции данных и приложений Цифрового предприятия. Обеспечивает единое информационное пространство на базе корпоративного портала, автоматизацию процессов обработки и ведения нормативно-справочной информации и включает следующие модули: система управления нормативно-справочной информацией; система управления коммуникациями и интеграцией информации; система управления контентом и электронными фондами на базе порталных технологий [1].

2. Программное обеспечение «Интеграционная платформа системы».

Обеспечивает посредством объединения средств интеграции данных и интеграции приложений создание единого информационного пространства, обеспечивающего своевременность и полноту обмена данными между интегрируемыми информационными системами (ИС) в режиме реального времени, а также единого функционального пространства, обеспечивающего доступ к функциям, реализуемым прикладным программным обеспечением отдельных интегрируемых ИС. Основными функциями программы являются: управление учетными данными пользователей; настройка и развертывание интеграционных сценариев; ведение репозитория интеграционных сценариев; мониторинг процесса выполнения интеграционных сценариев [2].

3. Комплекс программ управления ресурсами Цифрового предприятия.

Предназначен для эффективного управления ресурсами и деятельностью предприятия. Включает в себя модули, обеспечивающие расширенный диапазон функций, в том числе управление производственными программами, технологической базой, управление экономической деятельностью, бюджетным планированием, снабжением и сбытом, управление затратами, управление финансами, бухгалтерским и налоговым учетом, а также управление основными средствами, техническим обслуживанием и ремонтом оборудования [3].

4. Комплекс программ управление проектами и программами Цифрового предприятия.

Предназначен для обеспечения автоматизации процесса управления программами и проектами, обеспечения руководства предприятия, руководителей и команд управления проектами, методиками и средствами (инструментарием) управления проектами и программами на основе современных информационных технологий с учетом требований к эффективности, качеству и уровню информационной поддержки. Ориентирован на автоматизацию процессов управления НИОКР и производством [4].

5. Комплекс программ Управление производством Цифрового предприятия.

Предназначен для управления производством Цифрового предприятия на уровнях объемно-календарного, оперативного, внутрицехового планирования и диспетчирования. Включает в себя модули, обеспечивающие широкий диапазон функций, в том числе планирование и управление экспериментальным, опытным и серийным производством; управление планированием потребностей в материально-технических ресурсах; управление технологиями подготовки производства при изготовлении экспериментального образца изделий, при опытном и серийном производстве изделий; управление качеством в экспериментальном, опытном и серийном производстве; производственный документооборот [5].

6. Комплекс программных модулей для расширения функциональных возможностей PDM-системы.

Обеспечивает: автоматизированный расчёт исполнительных размеров мерительного инструмента (шаблоны, калибры гладкие) и его подбор; формирование информационно-удостоверяющего листа на отдельные документы или их группы; формирование групповой заявки (в виде обменного файла) от подразделения на добавление элементов в справочники нормативно-справочной информации и т.п. [6].

7. Программный модуль «Функционал дискреционного разграничения доступа».

Обеспечивает: выполнение требований СВТ-5 и СВТ-3 по классификации ФСТЭК России; диспетчирование доступа, используемое при обработке сведений ограниченного распространения, составляющих служебную или государственную тайну и обеспечивающее: идентификацию информационных объектов системы Лоцман: PLM, включая составные объекты, где содержательная часть хранится в файлах, а реквизитная часть хранится в записях БД; контроль доступа субъектов (пользователей системы Лоцман: PLM) к информационным объектам системы Лоцман: PLM в соответствии с матрицей доступа [7].

8. Программное обеспечение «Расчет показателей деятельности предприятия».

Обеспечивает представление результатов аналитической обработки информации о текущей эффективности предприятия с помощью показателей деятельности, позволяет проводить анализ дерева целей, исследовать динамическую OLAP-отчетность. Основными функциями программы являются: управление учетными данными пользователей; настройка алгоритмов расчета показателей; визуализация и расчет показателей; ведение иерархической структуры показателей; ведение дерева целей предприятия; ведение пользовательских пространств; представление динамической отчетности посредством OLAP-кубов [8].

9. Программное обеспечение «Портальные сервисы».

Обеспечивает функциональность корпоративной коммуникации, анализ и оценку характеристик управления предприятием, интеграцию разнородных данных и сервисов с использованием единого унифицированного доступа, а также доступ к данным и знаниям, электронной библиотеке. Основные задачи, решаемые программой: ведение новостных лент предприятия/подразделения; ведение календаря мероприятий предприятия/подразделения/сотрудника; управление заявками пользователей; управление блиц-опросами; реализация возможности online-обсуждения (форума); ведение электронной библиотеки предприятия; управление доской почета предприятия/подразделения; ведение списка вакансий; реализация личного кабинета сотрудника; предоставление функциональности электронной приемной руководства предприятия [9].

10. Библиотека программных кодов «Функции аналитической обработки информации».

Предназначена для автоматизированного получения актуальной информации из комплексной процессной модели посредством скриптов следующей информации: подсчет КПЭ, паспорт структурного подразделения, паспорт рабочего места, таксономическая модель, регламент и порядок процесса и т.п. Программа позволяет проводить структурированный анализ комплексной процессной модели предприятия [10].

11. Программное обеспечение СУ НСИ ТИС «РФЯЦ-ВНИИЭФ».

Предназначено для автоматизации процессов консолидации, первичной обработки и ведения нормативно-справочной информации (НСИ) и нормативно-технической документации (НТД).

12. Система автоматизированного нормирования труда - САНТ 3.0.

Позволяет нормировать различные виды работ: выполняемые на металлообрабатывающих станках (токарные, фрезерные и др.); выполняемые на специальном оборудовании, либо вручную (радиомонтажный комплекс, слесарно-сборочные, литейные и др.). Программа позволяет оперативно рассчитать: норму времени на выполнение технологической операции; комплексную норму времени (трудоемкость) на изготовление детали в соответствии с пооперационной (маршрутной) технологией (сумма норм времени на каждую операцию по изготовлению детали); трудоемкость выполненных работ за период времени с распределением трудоемкости по видам работ; стоимость работы по выполнению технологической операции и стоимость изготовления детали, изделия [11].

Комплексная процессная модель предприятия также представляет собой не абстрактное понятие, не статичный набор схем на бумажном носителе, а конкретный цифровой продукт – электронную базу данных интерактивных графических моделей деятельности предприятия (как основной, так и обеспечивающей) с различными атрибутами, свойствами и параметрами, сформированную с помощью специализированных сред моделирования бизнес-процессов [12,13,14].

Одна из первых успешных попыток внедрения концепции «Цифровое предприятие» в научно-производственную деятельность проведена в ряде предприятий ЯОК. На основе стандартов предприятия, научно-технической документации и

публикаций разработан прототип электронной модели предприятия ЯОК на основе электронной модели бизнес-процессов типового жизненного цикла изделий [15-19].

Рассмотренные программные комплексы в составе платформы «Цифровое предприятие» проходят успешную опытную эксплуатацию на одном из предприятий ЯОК и могут быть рекомендованы в качестве базового решения для построения цифровой информационно-технологической платформы предприятия.

Библиографический список

1. Костюков В.Е., Кривошеев О.В., Трищенко А.В., Жданов С.М., Тисков А.Н. ПО «Комплекс программ интегратор Цифрового предприятия», свид-во о гос. регистрации пр-мы для ЭВМ №2015662977 от 08.12.15.
2. Кривошеев О.В., Трищенко А.В., Жданов С.М., Цыгунька А.И., Чернышев А.Н., Гусельников Н.А. ПО «Интеграционная платформа Системы», свид-во о гос. регистрации пр-мы для ЭВМ №2014619456 от 16.09.2014.
3. Кривошеев О.В., Девяткин М.И., Жданов С.М., Милешин Г.А., Карташов Г.А., Тисков А.Н., Трищенко А.В. ПО «Комплекс программ управления ресурсами Цифрового предприятия V2», свид-во о гос. регистрации пр-мы для ЭВМ №2016619188 от 15.08.16.
4. Кривошеев О.В., Жигулин Д.В., Жданов С.М., Клименко А.А., Подвалов А.А., Филиппова И.С. ПО «Комплекс программ управление проектами и программами Цифрового предприятия», свид-во о гос. регистрации пр-мы для ЭВМ №2015662531 от 26.11.15.
5. Костюков В.Е., Кривошеев О.В., Пантелеева О.И., Жигачев В.Н., Хохлов А.В., Трищенко А.В., Занькова О.Н., Тихонов В.В. ПО «Комплекс программ управление производством Цифрового предприятия», свид-во о гос. регистрации пр-мы для ЭВМ №2016614921 от 12.05.16.
6. Костюков В.Е., Кривошеев О.В., Колодей Н.И., Захаров И.Н., Тихонов В.В., Сорокина И.В. ПО «Комплекс программных модулей для расширения функциональных возможностей PDM - системы», свид-во о гос. регистрации пр-мы для ЭВМ №2015661815 от 09.11.15.
7. Костюков В.Е., Кривошеев О.В., Будников В.И., Васильев Э.Н., Фиго Д.М. ПО «Программный модуль Функционал дискретного разграничения доступа», свид-во о гос. регистрации пр-мы для ЭВМ №2015661809 от 09.11.2015.
8. Кривошеев О.В., Трищенко А.В., Жданов С.М., Кондырин В.К., Симонов Д.Г., Каинов К.В. ПО «Расчет показателей деятельности предприятия», свид-во о гос. регистрации пр-мы для ЭВМ №2014618835 от 29.08.2014.
9. Кривошеев О.В., Трищенко А.В., Жданов С.М., Кондырин В.К., Симонов Д.Г., Зайцев К.В. ПО «Портальные сервисы», свид-во о гос. регистрации пр-мы для ЭВМ №2014618837 от 29.08.2014.
10. Кривошеев О.В., Карпенко С.В., Додонов Д.А. Библиотека программных кодов «Функции аналитической обработки информации», свид-во о гос. регистрации пр-мы для ЭВМ №2014615098 от 19.05.2014.
11. Система автоматизированного нормирования труда – САНТ 3.0, свид-во о гос. регистрации пр-мы для ЭВМ №2013660333 от 31.10.2013.
12. Костюков В.Е., Комаров Г.В., Свеженцев Г.В., Карташов Г.А., Соловьев В.П., Кривошеев О.В., Карпенко С.В. БД: Процессная модель РФЯЦ-ВНИИЭФ 2011, свид-во о гос. регистрации БД №2013620462 от 01.04.2013.
13. Костюков В.Е., Комаров Г.В., Свеженцев Г.В., Карташов Г.А., Соловьев В.П., Кривошеев О.В., Карпенко С.В. БД: Процессная модель РФЯЦ-ВНИИЭФ 2012 года «Как должно быть» свид-во о гос. регистрации БД №2013621192 от 19.09.2013.

14. Кривошеев О.В., Карпенко С.В.БД: Референтная модель предприятий ЯОК свид-во о гос. регистрации БД №2014620861 от 19.06.2014.
15. Пат. 2256053. Устройство управления электронными замками / Гончаров С.Н., Мартынов А.П., Николаев Д.Б., Фомченко В.Н. Заявл. 20.01.2003, опубл. 10.07.2005. Бюл. № 19.
16. О.В. Кривошеев, А.В. Трищенко, В.Л. Ведерников, Н.В. Горбатенко. Применение процессного подхода для повышения эффективности схемотехнического проектирования приборов радиоэлектронной аппаратуры. Сборник трудов XXXVII Всероссийской научно-технической конференции «Проблемы эффективности и безопасности функционирования сложных технических и информационных систем», г. Серпухов, 2018г., стр.37-43.
17. Пат. 2224075. Электронно-механический замок / Мартынов А.П., Николаев Д.Б., Сычев В.Т., Шишкин Г.И. Заявл. 28.11.2001, опубл. 20.02.2004. Бюл. № 5.
18. Пат. 2154722. Электронный кодовый замок / Мартынов А.П., Николаев Д.Б., Шишкин Г.И. Заявл. 27.08.1999, опубл. 20.08.2000. Бюл. № 23.
19. Ведерников В.Л., Горбатенко Н.В., Запонов Э.В., Кривошеев О.В., Николаев Д.Б. Проблемы внедрения процессного подхода в процессы разработки электронных систем радиоэлектронной аппаратуры XXXVII Всероссийская научно-техническая конференция «Проблемы эффективности и безопасности функционирования сложных технических и информационных систем», г. Серпухов, 2018г., стр. 31-37.

O.V. Krivosheev

**THE APPLICATION SOFTWARE SYSTEMS
AS A PART OF THE PLATFORM "DIGITAL ENTERPRISE"**

*«Russian Federal Nuclear Center – All-Russian Scientific Research
Institute of Experimental Physics» (RFNC-VNIIEF), Sarov*

The article deals with the application software complexes that are part of and operate within the framework of the information technology platform "Digital Enterprise".

Keywords: application software, software package, digital enterprise, information technology platform.

ПРИКЛАДНЫЕ ПРОГРАММНЫЕ КОМПЛЕКСЫ В СОСТАВЕ ПЛАТФОРМЫ «ЦИФРОВОЕ ПРЕДПРИЯТИЕ»

Российский федеральный ядерный центр – Всероссийский научно-исследовательский институт экспериментальной физики

В статье рассматриваются комплексы прикладного программного обеспечения, входящие в состав и функционирующие в рамках информационно-технологической платформы «Цифровое предприятие».

Ключевые слова: прикладное программное обеспечение, программный комплекс, цифровое предприятие, информационно-технологическая платформа.

Динамика производства и внешней среды все чаще и чаще приводит научно-производственные предприятия, выпускающие сложную наукоемкую продукцию, к пониманию необходимости внедрения концепции «Цифровое предприятие» - концепции комплексной поэтапной автоматизации современного предприятия.

Концепция (модель) «Цифровое предприятие» – это комплексная информационно-технологическая платформа, обеспечивающая информационную поддержку всех этапов жизненного цикла изделий научно-производственного предприятия, а также автоматизацию всего процесса создания изделия от управления требованиями до выпуска конечной продукции, ее гарантийного обслуживания и авторского надзора.

Электронная модель «Цифровое предприятие» может применяться как виртуальный (электронный) аналог реального предприятия для анализа возникающих проблем, проведения исследований и экспериментов для повышения качества продукции, сокращения сроков выполнения процессов и сокращения материальных затрат

Основа построения электронной цифровой модели предприятия – комплексная процессная модель предприятия, применяемая для наглядного и целостного представления деятельности конкретного предприятия, сформированная в ходе оптимизации как основных, так и вспомогательных, обеспечивающих бизнес-процессов предприятия, взаимосвязки моделей процессов основной и обеспечивающей деятельности с моделями организационной структуры, компетенциями, моделями целей, научно-производственной базы.

Комплексная информационно-технологическая платформа «Цифровое предприятие» с входящей в него системой промышленной автоматизации, основанная на комплексной процессной модели предприятия, формирует единое непрерывное сквозное информационное пространство ведения как научно-производственной, так и поддерживающей деятельности предприятия с едиными правилами управления жизненным циклом изделия, см. рис.1.

Однако платформа «Цифровое предприятие» это не просто «модное» направление исследования, не абстрактное понятие, а сложный многокомпонентный комплекс (система) взаимодействующих в едином информационном пространстве конкретных цифровых продуктов – прикладного программного обеспечения (прикладного ПО), баз данных, автоматизированных и информационных систем.

Помимо крупных автоматизированных информационно-программных систем предприятия (таких, как PDM и ERP системы, рассмотрение которых требует изложения в отдельной статье/докладе), в состав платформы «Цифровое предприятие» входят комплексы прикладного программного обеспечения, функционирующие в едином информационном пространстве ведения деятельности предприятия и обеспечивающие управление ресурсами предприятия, производством, безопасностью и т.п.



Рис. 1. Единое программное информационное пространство предприятия

Рассмотрим некоторые комплексы прикладного программного обеспечения, разработанные и функционирующие в рамках концепции «Цифровое предприятие» (местоположение рассматриваемых комплексов в составе платформа «Цифровое предприятие» приведено на рис.1).

1. Комплекс программ Интегратор Цифрового предприятия.

Предназначен для интеграции данных и приложений Цифрового предприятия. Обеспечивает единое информационное пространство на базе корпоративного портала, автоматизацию процессов обработки и ведения нормативно-справочной информации и включает следующие модули: система управления нормативно-справочной информацией; система управления коммуникациями и интеграцией информации; система управления контентом и электронными фондами на базе порталных технологий [1].

2. Программное обеспечение «Интеграционная платформа системы».

Обеспечивает посредством объединения средств интеграции данных и интеграции приложений создание единого информационного пространства, обеспечивающего своевременность и полноту обмена данными между интегрируемыми информационными системами (ИС) в режиме реального времени, а также единого функционального пространства, обеспечивающего доступ к функциям, реализуемым прикладным программным обеспечением отдельных интегрируемых ИС. Основными функциями программы являются: управление учетными данными пользователей; настройка и развертывание интеграционных сценариев; ведение репозитория интеграционных сценариев; мониторинг процесса выполнения интеграционных сценариев [2].

3. Комплекс программ управления ресурсами Цифрового предприятия.

Предназначен для эффективного управления ресурсами и деятельностью предприятия. Включает в себя модули, обеспечивающие расширенный диапазон функций, в том числе управление производственными программами, технологической базой, управление экономической деятельностью, бюджетным планированием, снабжением и сбытом, управление затратами, управление финансами, бухгалтерским и налоговым учетом, а также управление основными средствами, техническим обслуживанием и ремонтом оборудования [3].

4. Комплекс программ управление проектами и программами Цифрового предприятия.

Предназначен для обеспечения автоматизации процесса управления программами и проектами, обеспечения руководства предприятия, руководителей и команд управления проектами, методиками и средствами (инструментарием) управления проектами и программами на основе современных информационных технологий с учетом требований к эффективности, качеству и уровню информационной поддержки. Ориентирован на автоматизацию процессов управления НИОКР и производством [4].

5. Комплекс программ Управление производством Цифрового предприятия.

Предназначен для управления производством Цифрового предприятия на уровнях объемно-календарного, оперативного, внутрицехового планирования и диспетчирования. Включает в себя модули, обеспечивающие широкий диапазон функций, в том числе планирование и управление экспериментальным, опытным и серийным производством; управление планированием потребностей в материально-технических ресурсах; управление технологиями подготовки производства при изготовлении экспериментального образца изделий, при опытном и серийном производстве изделий; управление качеством в экспериментальном, опытном и серийном производстве; производственный документооборот [5].

6. Комплекс программных модулей для расширения функциональных возможностей PDM-системы.

Обеспечивает: автоматизированный расчёт исполнительных размеров мерительного инструмента (шаблоны, калибры гладкие) и его подбор; формирование информационно-удостоверяющего листа на отдельные документы или их группы; формирование групповой заявки (в виде обменного файла) от подразделения на добавление элементов в справочники нормативно-справочной информации и т.п. [6].

7. Программный модуль «Функционал дискреционного разграничения доступа».

Обеспечивает: выполнение требований СВТ-5 и СВТ-3 по классификации ФСТЭК России; диспетчирование доступа, используемое при обработке сведений ограниченного распространения, составляющих служебную или государственную тайну и обеспечивающее: идентификацию информационных объектов системы Лоцман: PLM, включая составные объекты, где содержательная часть хранится в файлах, а реквизитная часть хранится в записях БД; контроль доступа субъектов (пользователей системы Лоцман: PLM) к информационным объектам системы Лоцман: PLM в соответствии с матрицей доступа [7].

8. Программное обеспечение «Расчет показателей деятельности предприятия».

Обеспечивает представление результатов аналитической обработки информации о текущей эффективности предприятия с помощью показателей деятельности, позволяет проводить анализ дерева целей, исследовать динамическую OLAP-отчетность. Основными функциями программы являются: управление учетными данными пользователей; настройка алгоритмов расчета показателей; визуализация и расчет показателей; ведение иерархической структуры показателей; ведение дерева целей предприятия; ведение пользовательских пространств; представление динамической отчетности посредством OLAP-кубов [8].

9. Программное обеспечение «Портальные сервисы».

Обеспечивает функциональность корпоративной коммуникации, анализ и оценку характеристик управления предприятием, интеграцию разнородных данных и сервисов с использованием единого унифицированного доступа, а также доступ к данным и знаниям, электронной библиотеке. Основные задачи, решаемые программой: ведение новостных лент предприятия/подразделения; ведение календаря мероприятий предприятия/подразделения/сотрудника; управление заявками пользователей; управление блиц-опросами; реализация возможности online-обсуждения (форума); ведение электронной библиотеки предприятия; управление доской почета предприятия/подразделения; ведение списка вакансий; реализация личного кабинета сотрудника; предоставление функциональности электронной приемной руководства предприятия [9].

10. Библиотека программных кодов «Функции аналитической обработки информации».

Предназначена для автоматизированного получения актуальной информации из комплексной процессной модели посредством скриптов следующей информации: подсчет КПЭ, паспорт структурного подразделения, паспорт рабочего места, таксономическая

модель, регламент и порядок процесса и т.п. Программа позволяет проводить структурированный анализ комплексной процессной модели предприятия [10].

11. Программное обеспечение СУ НСИ ТИС «РФЯЦ-ВНИИЭФ».

Предназначено для автоматизации процессов консолидации, первичной обработки и ведения нормативно-справочной информации (НСИ) и нормативно-технической документации (НТД).

12. Система автоматизированного нормирования труда - САНТ 3.0.

Позволяет нормировать различные виды работ: выполняемые на металлообрабатывающих станках (токарные, фрезерные и др.); выполняемые на специальном оборудовании, либо вручную (радиомонтажный комплекс, слесарно-сборочные, литейные и др.). Программа позволяет оперативно рассчитать: норму времени на выполнение технологической операции; комплексную норму времени (трудоемкость) на изготовление детали в соответствии с пооперационной (маршрутной) технологией (сумма норм времени на каждую операцию по изготовлению детали); трудоемкость выполненных работ за период времени с распределением трудоемкости по видам работ; стоимость работы по выполнению технологической операции и стоимость изготовления детали, изделия [11].

Комплексная процессная модель предприятия также представляет собой не абстрактное понятие, не статичный набор схем на бумажном носителе, а конкретный цифровой продукт – электронную базу данных интерактивных графических моделей деятельности предприятия (как основной, так и обеспечивающей) с различными атрибутами, свойствами и параметрами, сформированную с помощью специализированных сред моделирования бизнес-процессов [12,13,14].

Одна из первых успешных попыток внедрения концепции «Цифровое предприятие» в научно-производственную деятельность проведена в ряде предприятий ЯОК. На основе стандартов предприятия, научно-технической документации и публикаций разработан прототип электронной модели предприятия ЯОК на основе электронной модели бизнес-процессов типового жизненного цикла изделий [15-19].

Рассмотренные программные комплексы в составе платформы «Цифровое предприятие» проходят успешную опытную эксплуатацию на одном из предприятий ЯОК и могут быть рекомендованы в качестве базового решения для построения цифровой информационно-технологической платформы предприятия.

Библиографический список

1. Костюков В.Е., Кривошеев О.В., Трищенко А.В., Жданов С.М., Тисков А.Н. ПО «Комплекс программ интегратор Цифрового предприятия», свид-во о гос. регистрации пр-мы для ЭВМ №2015662977 от 08.12.15.
2. Кривошеев О.В., Трищенко А.В., Жданов С.М., Цыгунька А.И., Чернышев А.Н., Гусельников Н.А. ПО «Интеграционная платформа Системы», свид-во о гос. регистрации пр-мы для ЭВМ №2014619456 от 16.09.2014.
3. Кривошеев О.В., Девяткин М.И., Жданов С.М., Милешин Г.А., Карташов Г.А., Тисков А.Н., Трищенко А.В. ПО «Комплекс программ управления ресурсами Цифрового предприятия V2», свид-во о гос. регистрации пр-мы для ЭВМ №2016619188 от 15.08.16.
4. Кривошеев О.В., Жигулин Д.В., Жданов С.М., Клименко А.А., Подвалов А.А., Филиппова И.С. ПО «Комплекс программ управление проектами и программами Цифрового предприятия», свид-во о гос. регистрации пр-мы для ЭВМ №2015662531 от 26.11.15.
5. Костюков В.Е., Кривошеев О.В., Пантелеева О.И., Жигачев В.Н., Хохлов А.В., Трищенко А.В., Занькова О.Н., Тихонов В.В. ПО «Комплекс программ управление производством Цифрового предприятия», свид-во о гос. регистрации пр-мы для ЭВМ №2016614921 от 12.05.16.

6. Костюков В.Е., Кривошеев О.В., Колодей Н.И., Захаров И.Н., Тихонов В.В., Сорокина И.В. ПО «Комплекс программных модулей для расширения функциональных возможностей PDM - системы», свид-во о гос. регистрации пр-мы для ЭВМ №2015661815 от 09.11.15.
7. Костюков В.Е., Кривошеев О.В., Будников В.И., Васильев Э.Н., Фиго Д.М. ПО «Программный модуль Функционал дискретного разграничения доступа», свид-во о гос. регистрации пр-мы для ЭВМ №2015661809 от 09.11.2015.
8. Кривошеев О.В., Трищенко А.В., Жданов С.М., Кондырин В.К., Симонов Д.Г., Каинов К.В. ПО «Расчет показателей деятельности предприятия», свид-во о гос. регистрации пр-мы для ЭВМ №2014618835 от 29.08.2014.
9. Кривошеев О.В., Трищенко А.В., Жданов С.М., Кондырин В.К., Симонов Д.Г., Зайцев К.В. ПО «Портальные сервисы», свид-во о гос. регистрации пр-мы для ЭВМ №2014618837 от 29.08.2014.
10. Кривошеев О.В., Карпенко С.В., Додонов Д.А. Библиотека программных кодов «Функции аналитической обработки информации», свид-во о гос. регистрации пр-мы для ЭВМ №2014615098 от 19.05.2014.
11. Система автоматизированного нормирования труда – САНТ 3.0, свид-во о гос. регистрации пр-мы для ЭВМ №2013660333 от 31.10.2013.
12. Костюков В.Е., Комаров Г.В., Свеженцев Г.В., Карташов Г.А., Соловьев В.П., Кривошеев О.В., Карпенко С.В. БД: Процессная модель РФЯЦ-ВНИИЭФ 2011, свид-во о гос. регистрации БД №2013620462 от 01.04.2013.
13. Костюков В.Е., Комаров Г.В., Свеженцев Г.В., Карташов Г.А., Соловьев В.П., Кривошеев О.В., Карпенко С.В. БД: Процессная модель РФЯЦ-ВНИИЭФ 2012 года «Как должно быть» свид-во о гос. регистрации БД №2013621192 от 19.09.2013.
14. Кривошеев О.В., Карпенко С.В. БД: Референтная модель предприятий ЯОК свид-во о гос. регистрации БД №2014620861 от 19.06.2014.
15. Пат. 2256053. Устройство управления электронными замками / Гончаров С.Н., Мартынов А.П., Николаев Д.Б., Фомченко В.Н. Заявл. 20.01.2003, опубл. 10.07.2005. Бюл. № 19.
16. О.В. Кривошеев, А.В. Трищенко, В.Л. Ведерников, Н.В. Горбатенко. Применение процессного подхода для повышения эффективности схемотехнического проектирования приборов радиоэлектронной аппаратуры. Сборник трудов XXXVII Всероссийской научно-технической конференции «Проблемы эффективности и безопасности функционирования сложных технических и информационных систем», г. Серпухов, 2018г., стр.37-43.
17. Пат. 2224075. Электронно-механический замок / Мартынов А.П., Николаев Д.Б., Сычев В.Т., Шишкин Г.И. Заявл. 28.11.2001, опубл. 20.02.2004. Бюл. № 5.
18. Пат. 2154722. Электронный кодовый замок / Мартынов А.П., Николаев Д.Б., Шишкин Г.И. Заявл. 27.08.1999, опубл. 20.08.2000. Бюл. № 23.
19. В.Л. Ведерников, Н.В. Горбатенко, Э.В. Запонов, О.В. Кривошеев, Д.Б. Николаев Проблемы внедрения процессного подхода в процессы разработки электронных систем радиоэлектронной аппаратуры XXXVII Всероссийская научно-техническая конференция «Проблемы эффективности и безопасности функционирования сложных технических и информационных систем», г. Серпухов, 2018г., стр. 31-37.

O.V. Krivosheev

**THE APPLICATION SOFTWARE SYSTEMS
AS A PART OF THE PLATFORM "DIGITAL ENTERPRISE"**

«Russian Federal Nuclear Center – All-Russian Scientific Research
Institute of Experimental Physics» (RFNC-VNIIEF), Sarov

The article deals with the application software complexes that are part of and operate within the framework of the information technology platform "Digital Enterprise".

Keywords: application software, software package, digital enterprise, information technology platform.

С.В. Глушков¹, А.С. Плехов², А.В. Соловьев¹, Ю.С. Федосенко¹

**КОНЦЕПЦИЯ ИНТЕЛЛЕКТУАЛИЗАЦИИ УПРАВЛЕНИЯ
ЭЛЕКТРОЭНЕРГЕТИЧЕСКОЙ СИСТЕМОЙ СУДНА
В КОНТЕКСТЕ ПАРАДИГМЫ БЕРЕЖЛИВОГО ПРОИЗВОДСТВА**

¹*Волжский государственный университет водного транспорта,*
²*Нижегородский государственный университет им. Н. И. Лобачевского*

Предлагается и обсуждается концепция энергосберегающего управления электроэнергетической системой судна. В основе концепции – подход, реализующий типовые сценарии функционирования судовых электропотребителей, основанный на данных online мониторинга состояния судового оборудования, эксплуатационной обстановки, параметров внешней среды, маршрутных и организационно-программных ограничений.

Ключевые слова: судовое оборудование, сценарии управления, ситуационные режимы, CASE-диаграммы

Развитию цифровых технологий, существенно расширяющих возможности автоматизации техники водного транспорта, во взаимодействии с парадигмой бережливого производства активно способствует интеллектуализация систем управления судовыми объектами [1].

Для электроэнергетических комплексов судов одно из перспективных направлений такой эволюции основано на реализации ситуационных режимов функционирования в рамках «библиотеки» сценариев, каждый из которых достаточно полно описывает технические и организационные особенности той или иной эксплуатационной ситуации.

В качестве примеров эксплуатационных ситуаций укажем режимы работы пассажирского судна, когда в целях экономии топлива, потребляемого дизель-генераторами, допустимо отключение большинства электропотребителей; в иных случаях может потребоваться запуск всех вспомогательных дизель-генераторов, обеспечивающих активацию аварийных систем; при движении судна на спокойной воде [2] основным является режим поддержание стабильного напряжения в сети.

При разработке вышеуказанных сценариев в целях повышения их адекватности необходимо учитывать мнения, рекомендации экспертов и специалистов, имеющих достаточный опыт эксплуатации судовых электроэнергетических комплексов. Такой подход позволяет создать систему управления, позволяющую снизить влияние человеческого фактора на функционирование всей электроэнергетической системы судна, повысить его экономическую эффективность и техническую безопасность [3].

Использование датчиков для мониторинга значений параметров судовых электропотребителей и ведение технического журнала в форме электронной базы данных позволит путем последующей специализированной обработки прогнозировать состояние основных рабочих агрегатов, упреждая их преждевременный выход из строя [4].

Таким образом, техническая реализация излагаемой концепции интеллектуализации управления электроэнергетической системой судна в контексте парадигмы бережливого производства позволит создать систему управления судовыми электростанциями нового поколения.

Для более детального представления концепции интеллектуализации управления электроэнергетической системой судна целесообразно считать её разделенной на группы, отражающие факторы, от которых зависит работа судового электрооборудования [5].

Для пассажирского судна это могут быть объекты, работающие в автоматическом режиме независимо от команд судоводителя или наличия пассажиров на борту.

Могут существовать группа объектов, зависящая от команд судоводителя и работающая в автоматизированном режиме, и группа объектов, функционирование которых непосредственно связано с наличием пассажиров на борту.

Каждой группе отвечает своя логическая цепочка обработки сигналов, что позволяет системе управления реагировать на наличие (отсутствие) пассажиров на борту, нагрузку электростанции и т.п.

Описанный подход позволяет в реальном масштабе времени не только назначать и даже автоматически определять режимы работы электрооборудования, но и конфигурировать функционирование судовых электростанций на вероятный набор электропотребителей, которые будут задействованы в соответствии со складывающейся обстановкой [6].

Визуализировать алгоритм, реализующий концепцию интеллектуального управления, удобно с помощью CASE-диаграмм [7]. Пример схемы типового блока CASE-диаграммы представлен на рис. 1. При этом этот блок может обладать несколькими каналами, каждый из которых независимо взаимодействует с внешней средой.

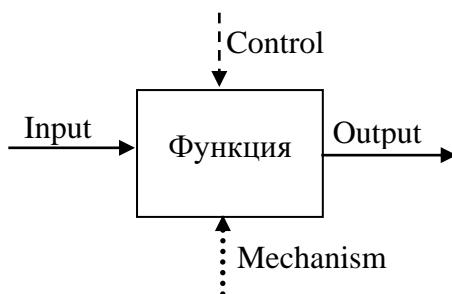


Рис. 1. Пример блока CASE-схемы

На вход (Input) поступает информация о ситуации или сигнал задания. Сигнал Control представляет собой правила, стратегии или стандарты работы оборудования, за которое отвечает автоматика. Вспомогательный сигнал (Mechanism) – обозначает сущности, выполняющие работу. Таковыми могут быть персонал, инструмент или специальное оборудование.

Выход (Output) – это результат всех операций, которые были выполнены в функциональном блоке.

Сигнал управления (Control) и вспомогательный сигнал (Mechanism) могут воздействовать как на отдельный блок, так и на группу блоков, если они логически объединены по принципу работы или по определенным путевым ситуациям или (и) на борту судна.

Примеры CASE-диаграмм алгоритма взаимодействия основных групп системы управления электропотреблением представлены на рис. 2, 3, 4.

Дальнейшее развитие обсуждаемой модели позволяет реализовать уровни приоритета потребления электроэнергии, с указанием какую нагрузку, в какое время допустимо отключить без негативных последствий для снижения энергопотребления, ускорить реакцию на управление электроэнергетической системой благодаря централизации управления электропитанием [8].

С применением централизованной системы расширяются возможности в части управления электрооборудованием. Например, подача команды в автоматизированном режиме на готовность электростанции к запуску аварийных систем в случае сценария работы «Максимальной технической безопасности судна в рейсе» при сильных штормовых условиях. Или наоборот, задействовать энергосберегающий режим в случае движения без пассажиров (или иной загрузки) в условиях спокойной путевой обстановки.

Технически такая система основывается на типовых цифровых технологиях и технических устройствах – главный контроллер или группа контроллеров обрабатывают информацию о техническом состоянии объектов судна, собранную на основе показаний датчиков [9].

Вместе с тем, результаты анализа потребляемого тока работающими и вновь подключающимися электропотребителями могут поступать в виде рекомендаций (вахтенному начальнику) о запуске дополнительного дизель-генератора или отключении уже действующей нагрузки для высвобождения необходимой мощности [10].



Рис. 2. CASE-диаграмма для объектов автоматического управления при регламенте номинальной нагрузки судовой электростанции

В заключение отметим, что предложенная в статье концепция интеллектуализации управления электроэнергетической системой судна в контексте парадигмы бережливого производства позволит существенно повысить оперативность управления электротехническими объектами судна и таким образом, сокращая объемы потребления топлива, будет способствовать снижению эксплуатационных расходов.

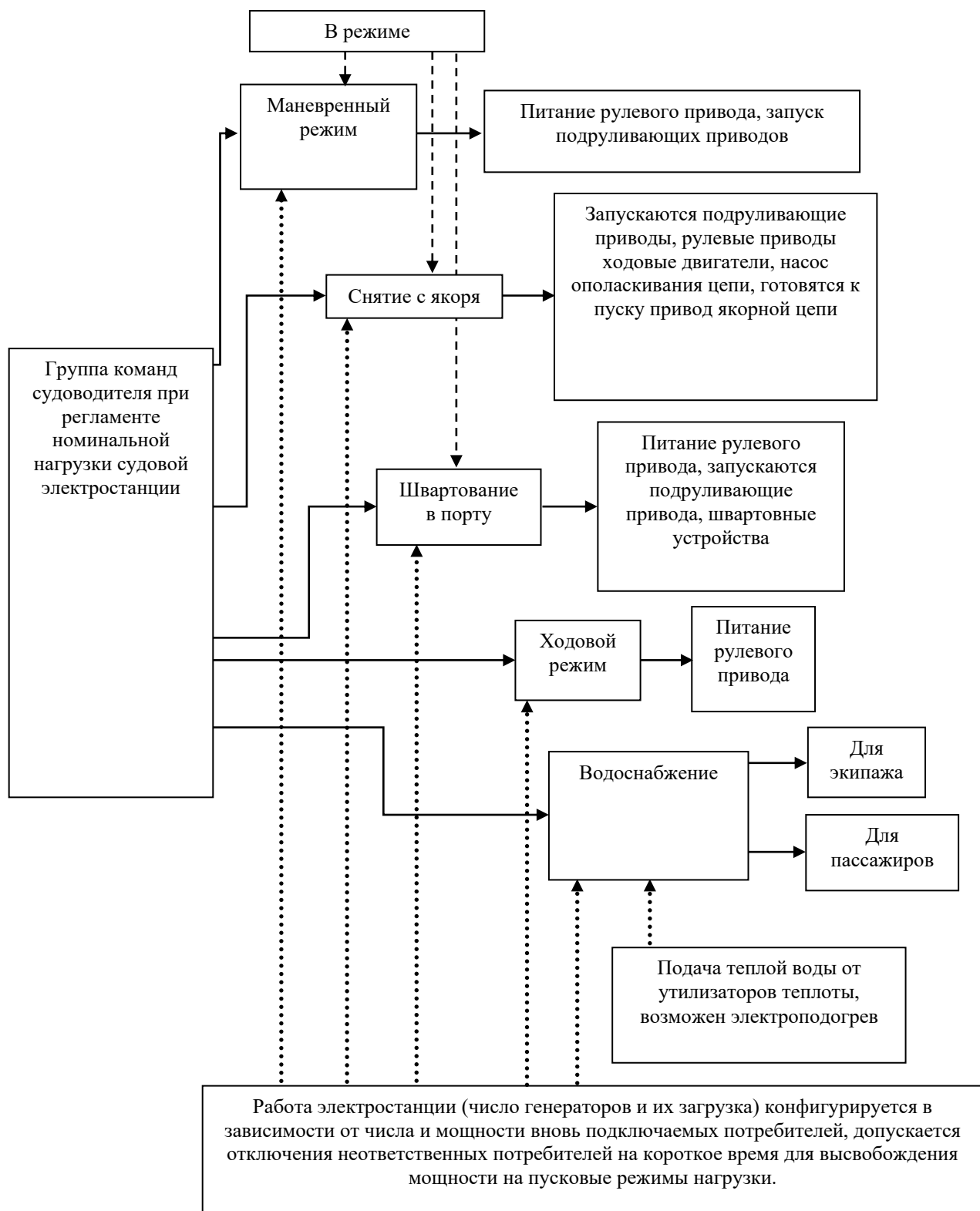


Рис. 3. CASE-диаграмма для объектов команд судоводителя при регламенте номинальной нагрузки судовой электростанции

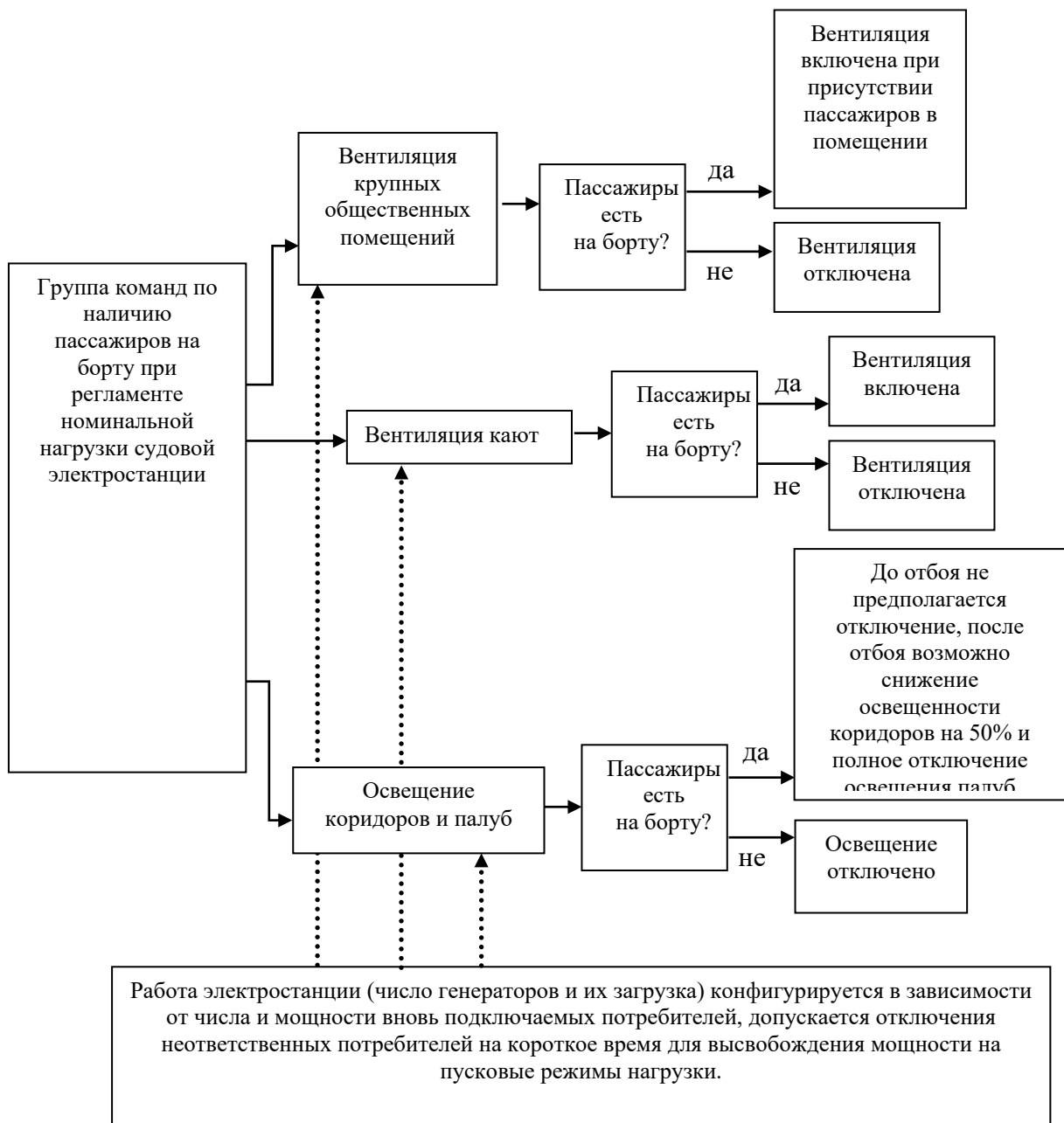


Рис. 4. CASE-диаграмма для объектов команд, зависящих от наличия пассажиров при регламенте номинальной нагрузки судовой электростанции

Библиографический список

1. Чуй С.А. Особенности внедрения бережливого производства. Переход к развитию производственной системы судостроения: отчет заседания / Комитет по судостроительной промышленности и морской технике. М.: 2016. – 59 с.
2. Соловьев А.В. Концепция единого целеориентированного управления судовой энергетической установкой // Вестник государственного университета морского и речного флота имени адмирала С.О Макарова. 2017. Т. 9. №5. С. 1027-1039.
4. Ляшенко В.Г. Управление судном. Влияние человеческого фактора на безопасность мореплавания: курс лекций / Херсон: Издательство ХГМА. 2016. – 168 с.
5. Мясников Ю.А. Методы прогнозирования технического состояния судового энергетического оборудования // Труды Крыловского государственного научного центра. 2018. Т. 368. №4. С. 117–132.

6. Коровин А.Г. Разработка методов влияния человеческого фактора на безопасность судна // Вестник Камчатского государственного технического университета. 2009. №10. С. 31-36.
7. Яцун А.С. Аппаратные средства цифрового управления мехатронными системами / Курск. 2012.
8. Карпычев В.Ю. Современные методологии и инструментальные средства поддержки проектирования информационных систем / Нижний Новгород: НГТУ им. Р.Е. Алексеева. 2014. – 97 с.
9. Губанов Ю. Электротехнические системы кораблей и судов: этапы развития, автоматизация. Концепция электрического корабля // Controlengineering Россия. 2017. №3(51). С. 24 – 27.
10. Буев С.А. Совершенствование технического аудита оборудования морских судов на основе системы термографического анализа: диссертация на соискание ученой степени кандидата технических наук / Мурманский ГТУ. 2014. – 152 с.
11. Дорри М.Х. Автоматизация управления морскими подвижными объектами // Проблемы управления. 2009. № 3.1. С. 94–102.

S.V. Glushkov¹, A.S. Plekhov², A.V. Solovev¹, Yu.S. Fedosenko¹

**THE CONCEPT OF INTELLECTUALIZATION CONTROL
SHIP'S ELECTRIC POWER SYSTEM
IN THE CONTEXT OF THE LEAN MANUFACTURING PARADIGM**

¹Volga state University of water transport, Nizhny Novgorod

²Lobachevsky state university of Nizhny Novgorod

The concept of energy-saving management of the ship's electric power system is proposed and discussed. The concept is based on an approach that implements typical scenarios for the operation of ship power consumers, based on data from on-line monitoring of the state of ship equipment, operational conditions, environmental parameters, route and organizational and program restrictions.

Keywords: ship equipment, control scenarios, situational modes, CASE diagrams

ПРОБЛЕМЫ ПРОЕКТНОГО МЕНЕДЖМЕНТА. ВЛИЯНИЕ ЕДИНОГО ИНФОРМАЦИОННОГО ПОЛЯ НА ВСЕХ ЭТАПАХ РУКОВОДСТВА ПРОЕКТОМ

¹*Волжский государственный университет водного транспорта,*

²*Филиал Федерального государственного унитарного предприятия «Российский федеральный ядерный центр – Всероссийский научно-исследовательский институт экспериментальной физики» «Научно-исследовательский институт измерительных систем им. Ю.Е. Седакова», г. Нижний Новгород*

Рассматривается на практике влияние единого информационного поля на эффективность управления проектом. На примере управления проектами изготовления и поставки технических средств для автоматизированных систем управления технологическими процессами проектов АЭС рассмотрены проблемы, с которыми сталкиваются разрабатывающие подразделения и руководитель проекта, и пути их решения.

Ключевые слова: проектный менеджмент, управление проектами, проект, цифровизация, базы данных, единое информационное поле.

Краткий экскурс по теме исследования

В разное время были свои концептуальные особенности, которые влияли как на методы и модели управления проектами на всех этапах жизненного цикла, так и на возможности совершенствования и оптимизации их. Исследуя принципы и методы управления проектами становится зримым настоящее положение дел в этой сфере и будущие перспективы, т.е. потенциальные направления развития данной управленческой сферы. Обобщая накопленный опыт при реализации крупных, в масштабе деятельности предприятия, проектов мы можем реконструировать процессы, благодаря которым проясняются скрытые факторы, объясняющие то, как глобализация, технический прогресс и цифровизация – репрезентированы в современной реальности проектного менеджмента. Не опираясь на опыт организации проектной деятельности, проблемы, с которыми сталкивались руководители проектов на разных этапах жизненного цикла одного и того же проекта, алгоритмы и модели для оптимизации проектного менеджмента, – невозможно разработать новое и универсальное программное обеспечение, которое объединяло бы информационное поле всех участников проекта, учитывая их потребности.

Наглядным примером оптимизации проектного менеджмента является разработка Графика исполнения договора (Далее График) с использованием программного обеспечения MS Project, структурированного по следующим основным этапам:

1. Заключение Договора;
 - 1.1. Заключение договоров с Субпоставщиками (при необходимости, если это подразумевается Договором);
2. Разработка и согласование Технических заданий и/или Технических условий на Оборудование, ТЗ и/или ТУ на единицы Оборудования;
3. Разработка РКД, ЭД;
4. Закупка;
5. Изготовление;
6. Испытания;
7. Упаковка, отгрузка и поставка Оборудования;
8. Начало Пусконаладочных работ.

Каждая фаза Графика должна содержать следующие данные:

- плановые даты начала и окончания выполнения фаз;
- связи с предшествующими и последующими фазами.

Помимо удобного программного обеспечения, для выполнения вышеперечисленных задач и разработки Графика используются наработки Генри Гантта и метод критического пути. Генри Гантт является прародителем управления проектами [1]. Диаграммы Гантта важны в истории современного управления проектами, потому что они признают преимущества разделения больших проектов на более мелкие управляемые задачи. Они также признают тот факт, что некоторые задачи могут зависеть друг от друга. Диаграммы Гантта до сих пор используются и считаются жизненно важным инструментом в инструментарии проектного менеджера.

Для полного погружения в исследования и разработку универсального продукта, необходимо разобраться в самом проектном менеджменте, ознакомившись с наибольшим количеством методов, алгоритмов и продуктов, и спроецировать его на выполняемые задачи в области автоматизированных систем управления технологическими процессами (Далее АСУ ТП).

Разработка, изготовление и поставка технических средств

Одним из этапов проектирования АЭС является разработка, изготовление и поставка технических средств (Далее ТС) для АСУ ТП. В данной статье мы рассмотрим процесс в рамках деятельности крупного предприятия ГК «Росатом», отдельные стадии и этапы которого представлены на рис. 1.

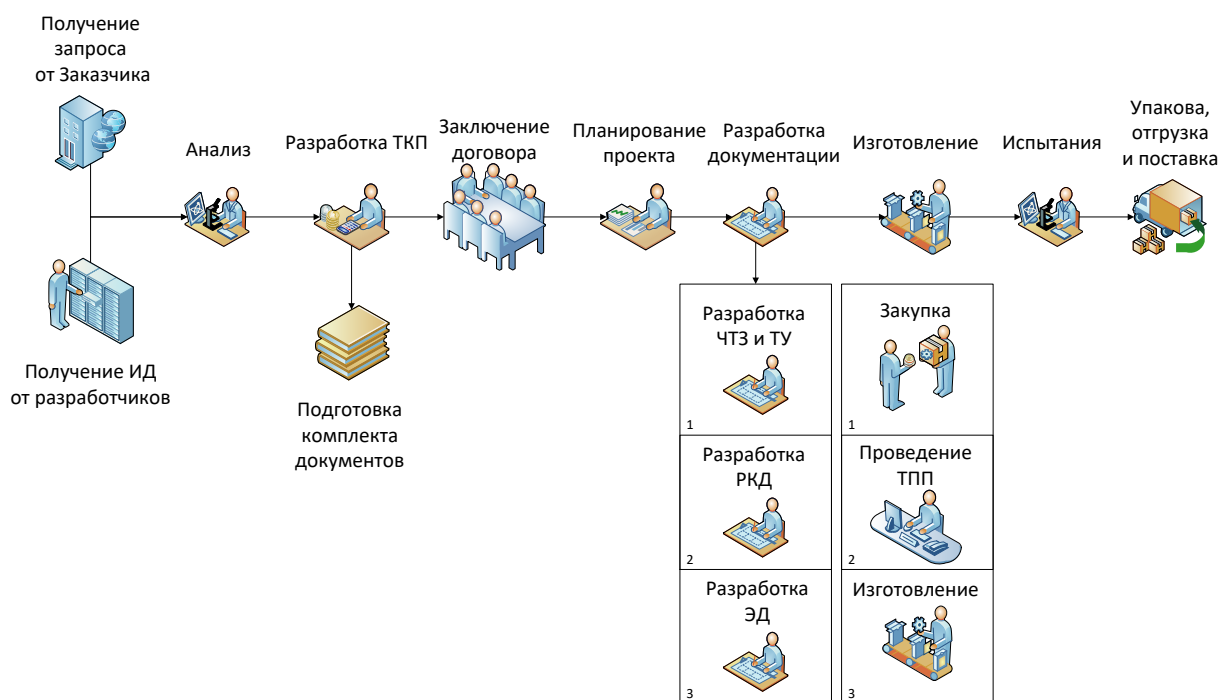


Рис. 1. Стадии и этапы изготовления и поставки ТС для АСУ ТП проектов АЭС.

Разработка технико-коммерческого предложения

Первая стадия в проектном менеджменте изготовления и поставки ТС, которая оказывает влияние на весь проект в целом, – определение исходных данных проекта, в том числе бюджет проекта, и разработка технико-коммерческого предложения (Далее ТКП). В настоящее время разработка ТКП производится по Единым отраслевым методическим указаниям по заполнению технико-коммерческих предложений на поставку компонентов АСУ ТП, введенным Приказом Госкорпорации «Росатом».

Основной задачей данной стадии является проведение наиболее точного расчета цены поставочной единицы оборудования в кратчайшие сроки. Стоит понимать, что чрезмерное превышение рыночной стоимости будет снижать конкурентоспособность предприятия, а заведомо низкая цена – делает проект неэффективным или убыточным.

Поставщик предоставляет Заказчику ТКП, содержащее данные для расчета НМЦ на компоненты АСУ ТП, заполненные по соответствующим формам, с подробными расшифровками затрат каждой статьи калькуляции и пояснительной запиской к расчету цены.

Для каждого проекта производится более 3-х десятков расчетов цен типовых изделий, каждый из которых имеет отличия в стоимости, в зависимости от требований проекта таких как:

- состав и характеристики изделий;
- состав работ, определенный условиями эксплуатации, требованиями Заказчика и применяемой нормативной базы;
- условия платежей;
- сроки реализации проекта;
- бюджет проекта.

Проблема 1. В процессе заполнения множества типовых форм под влиянием внешних факторов возникают случаи, когда требуется скорректировать общий параметр (накладные расходы, страховые сборы, прибыль, курс валюты, ставка НДС и др.) в большом количестве файлов с расчетами.

В связи с ограниченным количеством ресурсов и времени, выделяемых для разработки ТКП, собрать 100% объем исходных данных с участников процесса (разрабатывающих и конструкторских подразделений, финансово-экономических, логистических, договорных служб, производственных подразделений, службы снабжения и т.д.) разработки ТКП, а в случае необходимости оперативно внести коррекцию, в большинстве случаев не представляется возможным.

Этапы разработки технико-коммерческих предложений

В настоящее время на предприятии отсутствует унифицированный алгоритм взаимодействия участников процесса при разработке ТКП, соответствующий современным техническим возможностям, а существующий порядок действий разбивается на два этапа:

- На первом этапе ведется набор и обработка информации (входные данные). После получения запроса от Заказчика ответственный за предоставление ТКП получает исходные данные (далее – ИД) от разрабатывающих подразделений на бумажном носителе. Исходные данные содержат информацию об аналогах оборудования поставленного ранее по другому проекту, а также плановую трудоемкость на выполнение работ по проекту. Конфигурация аналогов технических средств (спецификация комплектующих и материалов) выгружается из электронного архива конструкторско-технологической документации (КТД), в котором отсутствует ценовая информация и ее источник. Отсутствие в архиве КТД ценовой информации обязывает разработчика ТКП использовать дополнительные источники (поиск в сети интернет, договора предшествующих периодов – база данных ERP-системы, ТКП изготовителей и поставщиков).

Проблема 2. Стоит обратить внимание, что конфигурация оборудования на этапе разработки ТКП формируется на основе аналогов уже реализованных проектов и в дальнейшем (в процессе реализации проекта) подвергается корректировкам.

- На втором этапе – осуществляется подготовка комплекта документов ТКП (выходные данные) по формам, соответствующим Единым отраслевым методическим указаниям. Разработка структур цен оборудования, сводной таблицы оборудования и спецификации оборудования осуществляется на основе расчетов, полученных на первом этапе.

Проблема 2. Отсутствие универсального программного продукта для разработки ТКП, вынуждает разработчика работать с информацией из разных массивов данных, объединяя и структурируя её по Единым отраслевым методическим указаниям по заполнению технико-коммерческих предложений на поставку компонентов АСУ ТП. Характерные человеческому фактору ошибки при ручном вводе большого количества однотипных данных в типовые формы сказываются на точности расчетов, а в случае обнаружения и исправления их – на быстродействии.

Для увеличения точности расчетов и оперативности предоставления ТКП необходимо совершенствование существующего процесса его разработки. Работа в едином информационном поле всех участников процесса создания продукции на предприятии позволит обеспечить сокращение сроков подготовки комплекта документов, снизить трудозатраты на разработку ТКП.

Заключение договора и планирование проекта

После заключения договора на предприятии назначается руководитель проекта (далее РП). Для успешного выполнения всех договорных обязательств РП необходима не только информация, полученная при разработке ТКП, но и дальнейшее управление изменениями (изменение уже имеющихся или появление новых переменных). На основе вышеперечисленного, руководитель проекта переходит к следующей стадии изготовления и поставки продукции – планирование проекта.

На стадии планирования руководитель проекта должен ознакомиться и проанализировать имеющуюся у него информацию и на её основе разработать сетевой план разработки, изготовления и поставки оборудования (Далее План). План подразумевает под собой последовательность детально декомпозированных задач по выполнению договорных обязательств, включающий в себя все этапы жизненного цикла производства.

Работы по сопровождению

После согласования и утверждения Плана, он направляется всем структурным подразделениям, участвующим в проекте. Все работы этих подразделений разбиваются на стадии: разработка РКД, закупка, изготовление, испытания, приёмка и отгрузка (шефмонтаж и шеф-наладка в случае необходимости). На протяжении всех этапов РП занимается работами по сопровождению, такими как управление закупкой, планирование платежей, планирование испытаний и др. Все эти действия приходится выполнять за рамками единого информационного поля. Отсутствие единых для всех форм и как следствие выполнение одних и тех же запросов разными людьми или подразделениями формирует разные типы данных.

Проблема 3. Таким образом, мы сталкиваемся с ограниченной и несовершенной системой, участники которой коммуницируют не находясь в рамках единого информационного поля унифицированного программного обеспечения.

Решение проблем и его достоинства

Предлагаемое решение выше поставленных проблем – разработка нового модуля в одной из существующих систем предприятия, который позволял бы одновременно участвовать в разработке ТКП всем участникам создания продукта. Использование унифицированного программного обеспечения увеличит точность и быстродействие разработки ТКП соответствующего отраслевым формам и создаст единое информационное поле.

Интегрирование в систему Плана на основе диаграммы Гантта и метода критического пути позволит демонстрировать объективную информацию о ходе работ проекта, включая расходование бюджета, состояние с заключением договоров и др. информацию.

Внедрение нового модуля влечет за собой синергию КД с закупкой, что многократно ускорит процедуру закупки. Исчезнет потребность в заведении огромного количества МТР за счет инициации закупочной кампании непосредственно из КД (формирование закупки на отдельные системы или на весь проект целиком). Появится возможность планирования закупочной кампании и платежей за поставку без участия разрабатывающих подразделений. Позволит на всех этапах ведения проекта получать полную и точную информацию по закупочной деятельности, визуализированную диаграммами. Будут видны изменения КД и в автоматическом режиме реализовано отображение дефицитов, на базе которых можно формировать заявки на закупку.

Важной особенностью единого информационного поля является определение и утверждение участников процесса в программной среде, что позволит перенести всю работу в дистанционный режим с предоставлением доступа для участия в выполнении общих задач. Внесение и коррекция уже существующих данных должна проводиться в соответствии с назначенными правами доступа. После внесения данных участником процесса, вся сопутствующая информация направляется на согласование в автоматическом режиме ответственному за процесс. После согласования ответственным за процесс, коррекция данных вносится в общую базу данных. В случае отказа согласования, данные возвращаются исполнителю для коррекции с пояснением причин.

Внедрение регулятивных принципов универсальной методологии управления проектами, подключая всех участников процесса, позволит создать единое информационное поле управления проектами с установкой перечня ограничений, что сократит сроки выполнения ряда процедур и исключит излишние действия.

Разработка программного обеспечения

Программное обеспечение разрабатывается в соответствии с V-моделью (Рис. 2): формируются требования, на следующем этапе формируется архитектура, дальше идет процесс проектирования (проект, проектирование, рабочее проектирование), затем стадии реализации, создания объекта. Далее идет обратно-восходящий процесс, когда нужно протестировать, провести различные процессы, связанные с верификацией, с приемкой, с проверкой и в конечном итоге сдать Заказчику, который должен убедиться, что он получил то, что задумал.

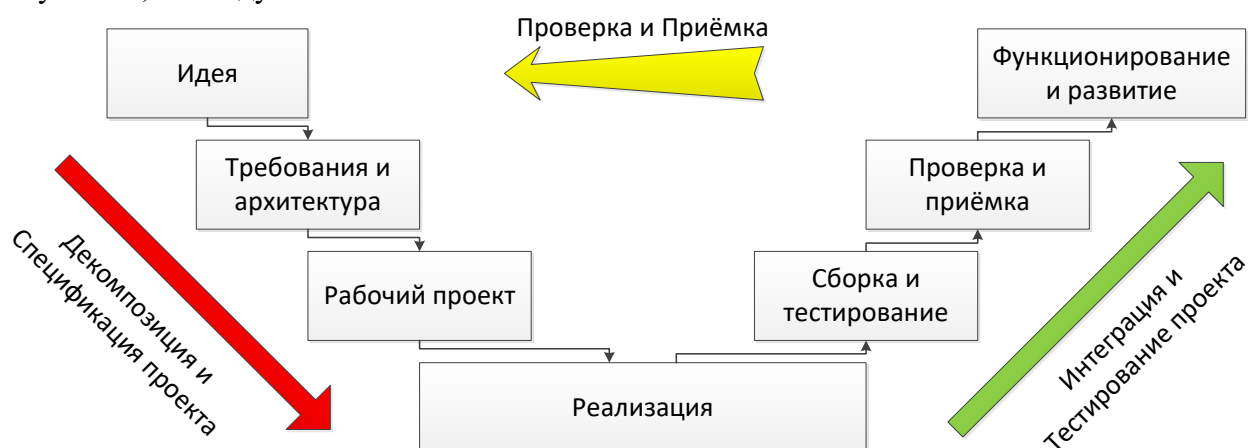


Рис. 2. V-модель разработки программного обеспечения.

Заключение

Информационные технологии, цифровизация и глобализация оказывают существенно влияние на развитие проектного менеджмента. Глобализация, ограниченность ресурсов, заинтересованные стороны, конкуренция, экономика и многие другие факторы способствуют преобразованию организаций и бизнес-среды [2]. В будущем, для достижения успеха, руководители проектов XXI века должны быть приспособлены к постоянным изменениям, неопределенностям, срывам и кризисам, а современные технологии и новые методологии управления проектами только способствуют достижению вышеупомянутых задач. Интеграция нового программного обеспечения является следующей ступенью в организации управления проектами в области, затрагиваемой в этой статье, таким образом, переход на нее неотвратимо произойдет в будущем. А вот когда произойдет это будущее, зависит только от нас.

Библиографический список

1. Chiu Y. (2010). Введение в историю управления проектами: с самых ранних времен до A.D.1900. Делфт: Эбурон.
2. Институт строительной промышленности. (2014). Навыки управления проектами будущего RS281-1. Строительство Индустриальный институт.

Artem D. Gluzdov², Andrey V. Kononov^{1,2}

PROBLEMS OF PROJECT MANAGEMENT. THE INFLUENCE OF A SINGLE INFORMATION FIELD AT THE STAGES OF PROJECT MANAGEMENT

¹Volga State University of Water Transport

²Branch of the Federal State Unitary Enterprise "Russian Federal Nuclear Center - All-Russian Scientific Research Institute of Experimental Physics" "Scientific Research Institute of Measuring Systems named after Yu.Ye. Sedakov", Nizhny Novgorod

In practice, the influence of a single information field on the effectiveness of project management is considered. On the example of project management for the manufacture and supply of technical means for automated control systems for technological processes of NPP projects, the problems faced by the developing departments and the project manager, and ways of solving them, are considered.

Keywords: project management, project, digitalization, data base, single information field.

ПРОБЛЕМЫ И ЗАДАЧИ РАЗВИТИЯ СИСТЕМ УПРАВЛЕНИЯ ОПЕРАЦИОННОЙ ПРОИЗВОДСТВЕННОЙ ДЕЯТЕЛЬНОСТЬЮ ВЫСОКОТЕХНОЛОГИЧНОГО ПРЕДПРИЯТИЯ

Филиал Федерального государственного унитарного предприятия «Российский федеральный ядерный центр – Всероссийский научно-исследовательский институт экспериментальной физики» «Научно-исследовательский институт измерительных систем им. Ю.Е. Седакова», г. Нижний Новгород

Суть исследования – выявить основные проблемы управления операционной деятельностью, выработать на практике механизм внедрения на предприятии диагностических, управляющих и моделирующих систем, исследовать на моделях работу интеллектуальных алгоритмов принятия решений в среде с неполной информацией. Для решения выявленных проблем ставится задача разработки программного обеспечения системы управления операционной производственной деятельности высокотехнологичного предприятия с возможностью перспективного управленческого планирования и анализа деятельности с учетом технологических особенностей, что в свою очередь позволит разрабатывать прогнозный сценарий развития операционной деятельности в будущем, на основе которого принимаются управленческие решения, снижающих риск их невыполнения.

Ключевые слова: система управления, мониторинг, процесс планирования работ в проекте.

Введение

Развитие современного высокотехнологичного предприятия – это постоянный рост его сложности, который проявляется как в расширении разнообразия технологий, использующих новые физические принципы и технические решения, так и в ускорении сменяемости и усложнении выпускаемой продукции. Сумма знаний, необходимых для эффективного проектирования и эксплуатации производства постоянно растет, что существенно увеличивает интеллектуальную нагрузку на проектировщика и производственника [1]. В частности, по мере внедрения новых форм автоматизации, базирующихся на электронной вычислительной технике, стремительно расширяется интеллектуальная составляющая труда персонала от управляющего до рабочего. Под управлением операционной производственной деятельностью (далее - ОД) понимаем целенаправленное функционирование производственного планирования для обеспечения равномерной и ритмичной работы предприятия по пути совершенствования рабочих процессов планирования, в результате осуществления которых лица принимающие решения (далее ЛПР) определяют, формулируют и количественно задают цели и пути их выполнения. Лидеры и их команды должны выходить на новый для осуществления контроля за ходом достижения целей предприятия, проверку управленческих решений и оценку их последствий. При управлении производственной деятельностью идет упор на аналитику учета, контроля, регулирования выполнения работ в режиме реального времени, также автоматизацию рабочих процессов [2]. В настоящее время отмечается тенденция роста потребности во внедрении информационных технологий, основанных на эффективных моделях [3], систем планирования операционных процессов (разработки и производства), сбора отраслевой статистики, сопровождения закупочной деятельности.

Проблемные факторы ОД

1. Недостаточная информационная база:

- не эффективные принципы, методы формирования базы данных для планирования;
- отсутствие необходимых средств, хранения, обработки и выдачи данных для их использования при решении плановых задач (объем услуг и производства продукции, трудовые и кадровые вопросы, материально-техническое обеспечение).

2. Степень влияния факторов риска на срывы операционной деятельности:

50% - длительность разработки и согласования требований на автоматизацию и техническое задание;

60% - длительная процедура выбора решений;

50% - проведение закупочных процедур;

40% - изменения требований к функционалу;

40% - потеря знаний о системе;

20% - недостаточность финансового обеспечения для реализации проекта и т.д. - усиление влияния человеческого фактора на производственную и ОД (влияние ЛПР).

Задачи, требующие решения

Выявленные проблемные факторы показывают актуальность решения следующих задач:

- разработки модели управления ОД предприятия с учетом динамичности рынка заказчиков;
- организации сбора, анализа, обобщения и распространения лучших практик управления;
- развитие и унификация средств информационного обмена на предприятии;
- актуализация и развитие системы сбора данных для мониторинга и контроля;
- создание информационной системы управления ОД предприятия, минимизирующую влияние человеческого фактора на всех стадиях жизненного цикла проекта.

Функциональность системы управления

Требуемую функциональность определим на основе анализа известного программного обеспечения в части планирования и управления работами. На рынке программного обеспечения в части планирования и управления работами существует множество программ. Рассмотрим две существующие программы по планированию на предприятии «Аванта» и «OPEN PLAN» - визуальный инструментарий представлен на рисунках 1 - 3.

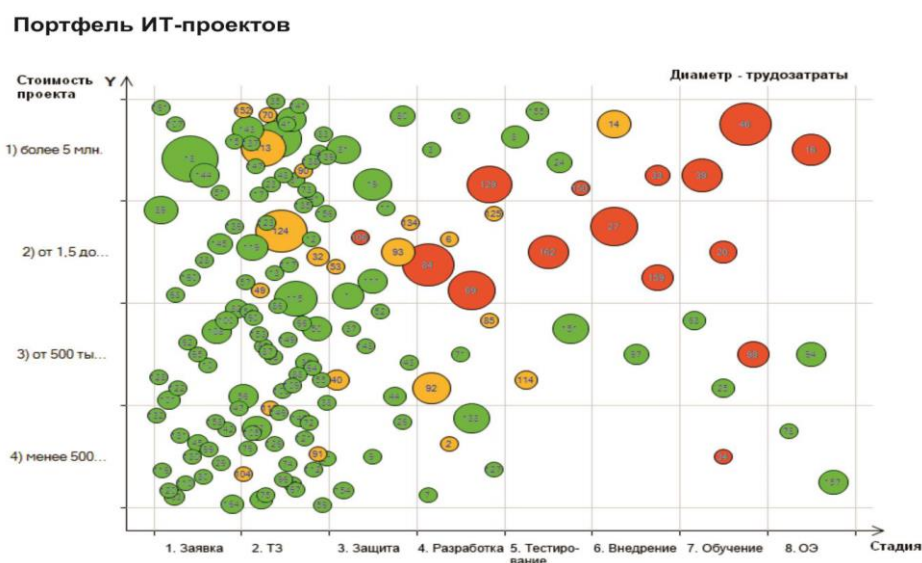


Рис. 1. Визуализация мониторинга портфеля ИТ - проектов

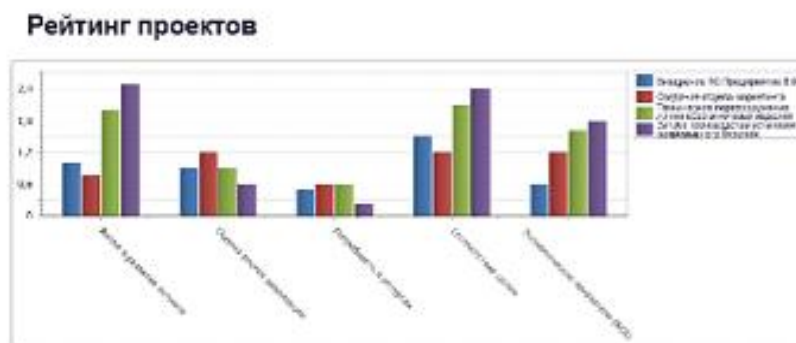


Рис. 2. Визуализация рейтинга проектов

Отклонение от бюджета

Проект	Общий Итог		
	Отклонение по стоимости	Индикатор	Тренд
Внедрение 1С:Предприятие 8.0	5 242.33	🔴	➡
Создание отдела маркетинга	3 887.20	🟡	⬇
Теорическое перевооружение линии сборки готовых изделий	1 874.24	🟢	⬆
Запуск производства установок непрерывного бурения	2 358.53	🟢	⬇

Рис. 3. Визуализация параметра мониторинга бюджета

Из рисунков видно, как выполняется визуализация данных. Так показывает рисунок 1, что информация выводится в виде раскрашенных кругов это проекты имеющие разную величину и стоимость, на рисунке 2 отображается рейтинг выполнения проектов (процент выполнения), рисунок 3 иллюстрирует контроль бюджета (норма или отклонение).

При пилотном применении данных программ на предприятии выявлены различные проблемы:

- цветовая палитра только в платной версии;
- ограниченная возможность детализации и планирования;
- ограниченный срок поддержки;
- ограниченная возможность крупномасштабного планирования;
- сужение функций, до функции календаря.

В данных программах при планировании работ один из способов мониторинга прохождения этапов жизненного цикла проекта (рисунок 4) базируется на построении диаграммы Ганта, применение которой имеет ряд недостатков:

- диаграмма Ганта не может проиллюстрировать несколько возможностей планирования в одном и том же графике;
- не позволяет прогнозировать ситуацию;
- отсутствие гибкости, в то время, как проекты не являются статичными (они постоянно меняются).

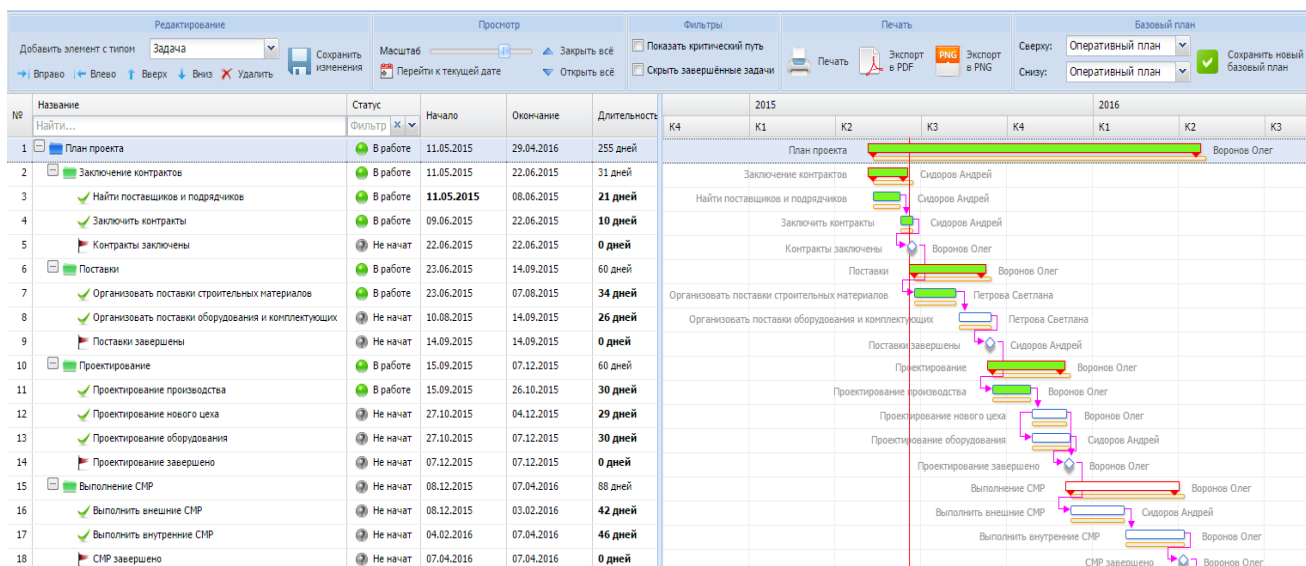


Рис. 4. Визуальный просмотр прохождения жизненного цикла проекта

Указанные проблемные факторы делают невозможным использование перечисленных программ планирования и управления операционной производственной деятельности высокотехнологичного предприятия с возможностью перспективного управленческого планирования. При разработке новых технологий управления операционной деятельностью необходимо разработать среду для моделирования поведения проектов под управлением интеллектуальных алгоритмов, что позволит сделать оболочки программ наглядными, легко воспринимаемыми и быстро решаемыми. В рамках такой технологии руководитель проекта:

- выбирает или формирует все составляющие «искусственного интеллекта» для разработки вариантов решений, средств и методов достижения поставленных целей;
- вводит в компьютер исходные данные, требуемые для работы «искусственного интеллекта»;
- анализирует приемлемость предложенных вариантов. Если приемлемых вариантов нет или их недостаточно, корректирует исходные данные, первоначально введенные в компьютер;
- выбирает наиболее приемлемый путь решения;
- по выбранному варианту рассчитываются и согласовываются с конкретными исполнителями средства, методы и время исполнения задания, ЛПР следит за исполнением заданий, разрабатывает варианты корректировок, выбирает наилучший и координирует процесс выполнения задания.

Функциональность разрабатываемой системы управления представлена в таблице 1, она сочетает диагностику, моделирование и управление рисками в проекте.

Таблица 1

Функциональность системы

Области проекта	Функциональность системы
Управление интеграцией	<ul style="list-style-type: none"> • Цветные статусы состояний для контроля показателей проекта (система светофора) • Автоматический расчет процента выполнения • Управление запросами на изменения • Управление правами доступа к проекту и к различной информации проекта • Ведение истории проекта

Управление сроками	<ul style="list-style-type: none"> • Определение длительностей работ • Определение условий и ограничений начала и окончания работ • Создание зависимостей между работами • Редактируемая диаграмма Гантта • Возможность экспорта/импорта плана проекта из/в MS Project • Автоматический расчет расписания методом критического пути • Возможность расчета предварительного и утверждения окончательного расписания • Уведомления и напоминания по отклонениям от заданного плана
Управление коммуникациями	<ul style="list-style-type: none"> • Формирование команды участников проекта • Средства совместной работы с документами проекта, управление версиями, согласования
Управление поставками	<ul style="list-style-type: none"> • Ведение базы контрактов проекта в разрезе контрагентов, статей бюджета • Ведение списка поставляемого оборудования • Контроль сроков поставок
Управление стоимостью	<ul style="list-style-type: none"> • Создание платежного календаря, в котором даты платежей могут быть как фиксированными, так и связанными с датами задач проекта • Прогноз превышений бюджета по освоенному объему • Финансовые отчеты с разрезом по статьям затрат • Утверждение бюджета с защитой от редактирования • Планирование бюджета
Управление поставками	<ul style="list-style-type: none"> • Контроль сроков оплат

Заключение

Определены «узкие места» существующих систем управления операционной деятельностью. Сформулированы задачи развития на основе автоматизации процедуры принятия решений и мониторинга текущей операционной деятельности. Такая система управления будет являться ежедневным рабочим инструментом пользователя, с помощью которого координируются и реализуются текущие бизнес задачи. Это позволит ЛПР в удобной среде комфортно генерировать сбор, обработку и оценку информации о результатах деятельности подразделений, составление отчетности, выявление отклонений от установленных промежуточных целей и причин этих отклонений, а также осуществить определение действий, необходимых для достижения конечной цели.

Библиографический список

1. Будыльский А.В. Управление ресурсами проекта на основе агентных технологий: дис. канд. техн. наук: 05.13.10/ Будыльский Александр Викторович. – Астрахань, 2014. – 145 с.
2. Макаров А.А., Ничипорук А.О. Планирование и организация новых производств, Нижний Новгород, 2009. - 20 с.
3. Ивлев М.А. Методология и технологии управления социально-экономическими системами при проектировании и развитии инновационного продукта: Диссертация докт. техн. наук: 05.13.10 / Уфимский государственный авиационный технический университет. 2014. - 32 с.

D.E. Ryabov

**PROBLEMS AND TASKS OF DEVELOPMENT OF MANAGEMENT SYSTEMS
FOR OPERATIONAL PRODUCTION ACTIVITY OF THE HIGH-TECH ENTERPRISE**

Branch of the Federal State Unitary Enterprise "Russian Federal Nuclear Center - All-Russian Scientific Research Institute of Experimental Physics" "Scientific Research Institute of Measuring Systems named after Yu.Ye. Sedakov", Nizhny Novgorod

The essence of the study is to identify the main problems of operating activities management, to develop in practice a mechanism for the implementation of diagnostic, control and modeling systems at the enterprise, to investigate the operation of intelligent decision-making algorithms in an environment with incomplete information on models. To solve the identified problems, the task is to develop software for a management system for operational production activities of a high-tech enterprise with the possibility of long-term management planning and analysis of activities, taking into account technological features, which in turn will allow developing a predictive scenario for the development of operational activities in the future, on the basis of which management decisions are made, reducing the risk of their failure.

Keywords: process of control system, monitoring, the planning work in the project.

К ВОПРОСУ ОПРЕДЕЛЕНИЯ ПОРОГОВЫХ ЗНАЧЕНИЙ ИЗМЕНЕНИЯ ХАРАКТЕРИСТИК ОСТОЙЧИВОСТИ СУДНА

¹Волжский государственный университет водного транспорта,
²ООО "Научно-производственная фирма "Штурвал", г. Нижний Новгород

Рассматривается алгоритм определения пороговых значений изменения характеристик остойчивости. Алгоритм разработан для использования в программной реализации информационной цифровой системы поддержки принятия решений об использовании капитаном штатных средств спасения при угрозе затопления и опрокидывания речного водоизмещающего судна. Приведены результаты расчетов гидростатических характеристик корпуса и посадки судна в случае затопления двух отсеков и действии наклоняющего момента для корпуса теплохода массового проекта 26-37.

Ключевые слова: кривые интенсивности изменения характеристик остойчивости, пороговые значения изменения характеристик остойчивости, гидростатические характеристики корпуса судна, посадка судна, диаграмма статической остойчивости

Для оценки опасности возникновения опрокидывания или заливания судна предлагается выполнять анализ кривых интенсивности изменения характеристик остойчивости [1] с течением времени. Кривые строятся по результатам определения пороговых значений характеристик остойчивости, вычисляемых с некоторым шагом по времени в процессе эксплуатации судна. В качестве пороговых значений приняты точки на диаграммах статической остойчивости, определяющие окончание прямолинейного участка восходящей ветви (точка «А») и максимум (точка «В») [2].

Далее представлен алгоритм определения пороговых значений характеристик остойчивости на примере судна с параметрами речного водоизмещающего пассажирского теплохода массового проекта 26-37 и построения кривых интенсивности их изменения.

Теоретически точка «А» определяет окончание прямолинейного участка восходящей ветви ДСО [3] и может быть определена по графику первой производной ДСО, как окончание первого горизонтального участка графика первой производной. Однако, как показывают исследования динамики и анализ плавучести и остойчивости судна [4], находящегося в эксплуатации, у речных водоизмещающих судов восходящая ветвь математически не всегда линейная. В связи с этим точку «А» предлагается определять путем исследования первой производной ДСО следующим образом: продифференцировать диаграмму по углу крена [5], найти относительные изменения последующих значений к предыдущим на каждом шаге, проверить условие об отклонении значений относительных величин более чем на 5% как свидетельство о прохождении точки «А» на предыдущем шаге.

Данное условие в алгоритме записывается следующим образом:

$$\text{if } \left| \frac{h_{\theta+1}}{h_{\theta}} - 1 \right| > 0,05: \quad (1)$$
$$h_{\theta-1} \rightarrow A$$

Точка «В», как наибольшая ордината ДСО [6], соответствует предельному кренящему моменту, статическое приложение которого еще не вызывает опрокидывания судна. После прохождения этой точки начинается нисходящая ветвь ДСО, соответствующая зоне неустойчивого равновесия судна, т.е. фактической потере остойчивости при статическом приложении кренящего момента. «В» – критическая точка

ДСО, где производные первого порядка обращаются в ноль. Геометрически это означает, что касательная к графику функции горизонтальна.

Данное условие записывается следующим образом:

$$\text{if } l_{\theta} = l_{max}: \quad (2)$$

$$l_{\theta} \rightarrow B$$

или

$$\text{if } h_{\theta+1} < 0: \quad (3)$$

$$h_{\theta} \rightarrow B$$

Описанная операция производится дискретно с некоторым шагом по времени. Для каждого шага отстраиваются ДСО, определяются пороговые значения (точки «А» и «В»). Таким образом получается ансамбль ДСО с определёнными пороговыми значениями. Отслеживая изменения пороговых значений для каждой последующей ДСО, можно получить графические кривые изменения пороговых значений (рис.2). Перестраивая эти кривые в функции от времени (для которого построена каждая ДСО), получаем кривые интенсивности изменения пороговых значений характеристик остойчивости судна. В таблице 1 представлен пример нахождения точки «А» для ДСО, соответствующей затоплению отсека до уровня 1 м.

Таблица 1

Пример нахождения точки «А»

θ , град.	h_{θ} , м	$\left \frac{h_{\theta+1}}{h_{\theta}} - 1 \right $	θ , град.	h_{θ} , м	$\left \frac{h_{\theta+1}}{h_{\theta}} - 1 \right $
1	0.02595	0.000	25	-0.00665	2.023
2	0.0262	0.010	26	-0.01015	0.526
3	0.0266	0.015	27	-0.0139	0.369
4	0.02695	0.013	28	-0.01745	0.255
5	0.02745	0.019	29	-0.02075	0.189
6	0.0282	0.027	30	-0.0262	0.263
7	0.029	0.028	31	-0.0286	0.092
8	0.02975	0.026	32	-0.031	0.084
9	0.03045	0.024	33	-0.03335	0.076
10	0.03105	0.020	34	-0.03565	0.069
11	0.0316	0.018	35	-0.03785	0.062
12	0.03215	0.017	36	-0.0395	0.044
13	0.03255	0.012	37	-0.04005	0.014
14	0.0329	0.011	38	-0.03985	0.005
15	0.0331	0.006	39	-0.0394	0.011
16	0.03295	0.005	40	-0.03885	0.014
17	0.0321	0.026	41	-0.0383	0.014
18	0.02915	0.092	42	-0.03775	0.014
19	0.0242	0.170	43	-0.0372	0.015
20	0.01865	0.229	44	-0.03665	0.015
21	0.0121	0.351	45	-0.03615	0.014
22	0.0073	0.397	46	-0.0357	0.012
23	0.0021	0.712	47	-0.03525	0.013
24	-0.0022	2.048	48	-0.0348	0.013

Для иллюстрации работы алгоритма рассмотрен пример оценки изменения устойчивости теплохода пр. 26-37 при наличии внешней нагрузки и затоплении отсека. Представлены результаты расчетов, в которых в качестве внешней нагрузки приняты: момент, кренящий $M_{кр} = 1500$ кН·м и момент дифференцирующий $M_{диф} = -9000$ кН·м. Произведён расчет гидростатических характеристик и построение ансамбля диаграмм статической устойчивости (ДСО) для последовательных состояний судна [7], соответствующих началу поступления воды в отсеки и увеличению её уровня в затопляемых отсеках с шагом 0,5 м.

Из таблицы 1 видно, что на основании выполнения условия (1) точка «А» соответствует углу крена 17° (рис. 1).

В таблице 2 представлен пример нахождения точки «В» для ДСО, соответствующей тому же уровню воды в отсеке 1 м.

Таблица 2

Пример нахождения точки «В»

θ , град.	l_θ , м	h_θ , м	θ , град.	l_θ , м	h_θ , м
1	-0.072	0.02595	25	0.5141	-0.00665
2	-0.046	0.0262	26	0.5032	-0.01015
3	-0.0196	0.0266	27	0.4938	-0.0139
4	0.0072	0.02695	28	0.4754	-0.01745
5	0.0343	0.02745	29	0.4589	-0.02075
6	0.0621	0.0282	30	0.4339	-0.0262
7	0.0907	0.029	31	0.4065	-0.0286
8	0.1201	0.02975	32	0.3767	-0.031
9	0.1502	0.03045	33	0.3445	-0.03335
10	0.181	0.03105	34	0.31	-0.03565
11	0.2123	0.0316	35	0.2732	-0.03785
12	0.2442	0.03215	36	0.2343	-0.0395
13	0.2766	0.03255	37	0.1942	-0.04005
14	0.3093	0.0329	38	0.1542	-0.03985
15	0.3424	0.0331	39	0.1145	-0.0394
16	0.3755	0.03295	40	0.0754	-0.03885
17	0.4083	0.0321	41	0.0368	-0.0383
18	0.4397	0.02915	42	-0.0012	-0.03775
19	0.4666	0.0242	43	-0.0387	-0.0372
20	0.4881	0.01865	44	-0.0756	-0.03665
21	0.5039	0.0121	45	-0.112	-0.03615
22	0.5123	0.0073	46	-0.1479	-0.0357
23	0.5185	0.0021	47	-0.1834	-0.03525
24	0.5165	-0.0022	48	-0.2184	-0.0348

Из таблицы 2 видно, что на основании выполнения условия (3) точка «В» соответствует углу крена 23° (рис. 1)

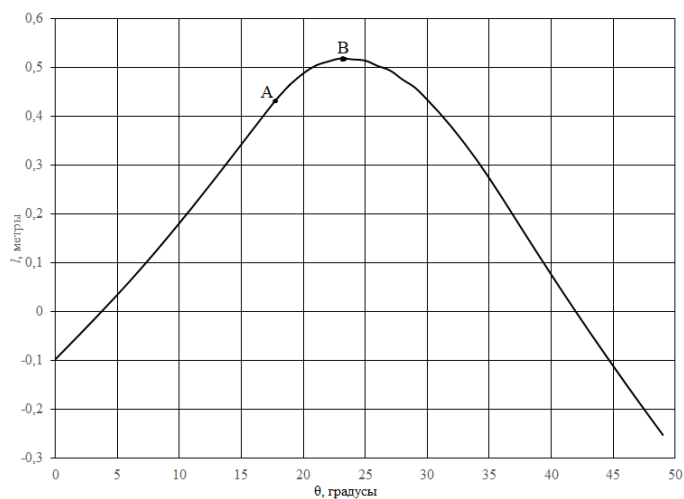


Рис. 1. Пример нахождения точек «А» и «В».

На ансамбле ДСО (рис. 2) на каждой диаграмме отмечены точки «А» и «В». Точки «А» соединены и образуют кривую первого порогового значения изменения характеристик остойчивости судна $\alpha(\theta)$; «В» - кривую второго порогового значения изменения характеристик остойчивости судна $\beta(\theta)$.

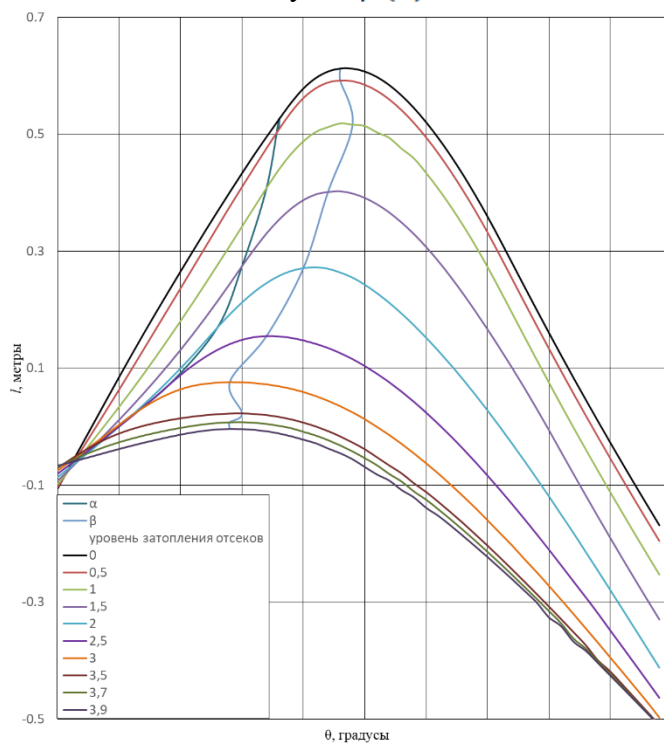


Рис. 2. Кривые пороговых значений изменения характеристик остойчивости судна

В программе ДИНАМИКА [8] был рассмотрен процесс изменения посадки судна с течением времени t , с (рис. 3). В результате мгновенного приложения моментов ($M_{кр}=1500$ кН·м, $M_{диф}=-9000$ кН·м.) судно в течение 90 секунд переходит к плаванию в положении равновесия с креном $\theta = 2,8^\circ$ на правый борт и дифферентом $\psi = 0,1^\circ$ корму. Затем, моделируя появление пробоины, к этой нагрузке добавляется постепенное затопление двух кормовых отсеков со скоростью поступления воды $0,95$ м³/мин. При этом контролируется уровень затопления отсеков $T_{зат}$, м.

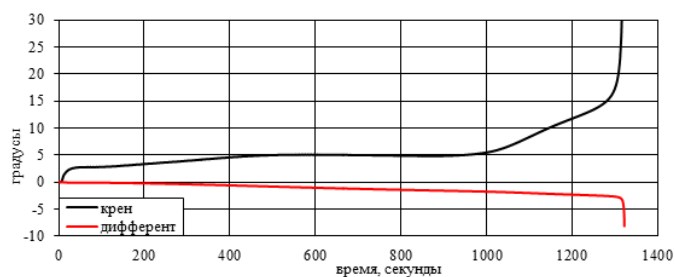


Рис. 3. Моделирование динамики посадки судна

Путем сопоставления данных исследования динамики гибели судна с результатами построения кривых пороговых значений изменения характеристик остойчивости от времени $\alpha(t)$ и $\beta(t)$ (таблица 3) установлено, что после достижения фактическим углом крена крайнего значения кривой $\alpha(t)$, наблюдается резкий рост угла крена судна (рис.3). Достижение характеристиками остойчивости значения, соответствующего точке «А» (рис. 4) является предпосылкой приближения к началу скоротечной аварии или к конечной стадии продолжительной аварии [9].

Момент достижения значения кривой второго порогового значения изменения характеристик остойчивости $\beta(t)$ величины фактического угла крена является моментом гибели судна (рис. 4).

Таблица 3

Результаты исследования

Т _{заг} , м	«А», град.	«В», град.	θ , град.	Время, с
0	18	23	2,8	90
0,5	18	23	3,1	166
1	17	24	3,7	270
1,5	15	22	4,5	385
2	13	20	5,0	519
2,5	9	17	5,0	673
3	5	14	4,8	858
3,5	-	15	7,9	1086
3,7	-	14	11,7	1197
3,9	7	14		1320

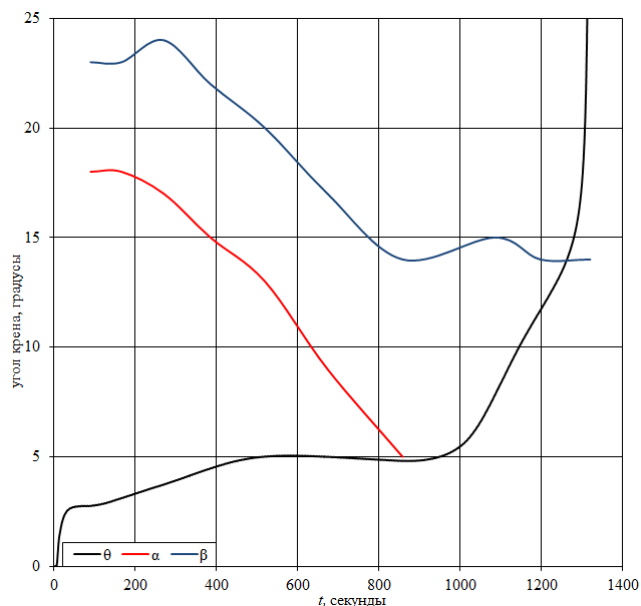


Рис. 4. Результаты исследования

По результатам проведенного анализа можно сделать вывод о том, что кривая первого порогового значения изменения характеристик остойчивости $\alpha(t)$ является искомым параметром, характеризующим остойчивость судна до уровня о необходимости использования спасательных средств.

Изменение интенсивности кривой второго порогового значения изменения характеристик остойчивости $\beta(t)$ (переход от снижения к росту) так же является параметром, характеризующим остойчивость судна до уровня необходимости использования спасательных средств.

Работа выполнена при поддержке гранта «Фонда содействия развитию малых форм предприятий в научно-технической сфере» (код договора 0065581).

Библиографический список

1. Валяев А.В., Федосенко Ю.С., Лукина Е.А. Моделирование динамики судна для системы поддержки принятия решений при угрозе скоротечного затопления /13-я мультиконференция по проблемам управления, материалы конференции «Управление в морских системах» (УМС-2020) 7–8 октября Санкт-Петербург, 2020. –с.124-126.
2. Валяев А.В., Лукина Е.А., Любимов В.И., Федосенко Ю.С. Мониторинг остойчивости в системе поддержки принятия решений об использовании средств спасения при угрозе скоротечного затопления речного судна / Морские интеллектуальные технологии. 2018. Т. 2. №4(42). С. 18-24.
3. Алферьев М.Я. Теория корабля. Плавуемость, остойчивость, непотопляемость и спуск судов на воду - М.: Транспорт, 1972. - 448 с.
4. Валяев А.В., Лукина Е.А., Федосенко Ю.С. Анализ диаграмм статической остойчивости в автоматизированной системе поддержки принятия решений об использовании судовых спасательных средств / Труды 22-го международного научно-промышленного форума «Великие реки». Материалы научно-методической конференции профессорско-преподавательского состава, аспирантов, специалистов и студентов «Проблемы использования и инновационного развития внутренних водных путей в бассейнах великих рек». – Нижний Новгород: ФГБОУ ВО «ВГУВТ», 2020. – http://vf-река-море.pf/2020/PDF/8_3.pdf (дата обращения 05.02.2021)
5. Валяев А.В., Лукина Е.А. Упреждающий мониторинг аварийной ситуации речного водоизмещающего судна / Информационные системы и технологии. 2020: [Электронный ресурс]: Сб. материалов XXV Международной научно-технической конференции. – Н. Новгород: НГТУ им. Р.Е. Алексеева. 2019. – CD диск.
6. Милова И.И. Теория корабля. Часть I. Статика судов внутреннего и смешанного плавания: учебное пособие для студентов специальности 180101 «Кораблестроение» / И.И. Милова. - Нижний Новгород: Издательство ФГОУ ВПО «ВГАВТ». 2008
7. Ваганов А.Б. Краснокутский И.Д. Свидетельство о государственной Регистрации программы для ЭВМ, №2014612390 «Расчет гидростатических характеристик и диаграмм остойчивости судовых корпусов сложной геометрической формы при произвольной посадке».
8. Ваганов А.Б., Краснокутский И.Д. Свидетельство о государственной Регистрации программы для ЭВМ, №2014612348 «Расчет качки судов и плавучих сооружений на морском волнении».
9. Александров М.Н. Безопасность человека на море / М.Н. Александров. - Л.: «Судостроение». 1983. – 209 с.

Aleksandr V. Valyaev^{1,2}, Aleksey G. Kljuhev², Evgenia A. Lukina^{1,2}, Jury S. Fedosenko¹

**TOUCHING THE DETERMINATION OF THE THRESHOLD VALUE CHANGES
IN CHARACTERISTICS OF THE VESSEL'S STABILITY**

*¹Volga State University of Water Transport
²LLC "Research and production firm "Shturval"*

An algorithm for determining threshold values of changes in stability characteristics is considered. The algorithm was developed for use in the software implementation of an information digital system for supporting decision-making on the use of standard rescue equipment by the captain in the event of a threat of flooding and overturning of a river displacement vessel. The results of calculations of the hydrostatic characteristics of the hull and landing of the vessel in the case of flooding of two compartments and the action of the inclining moment for the hull of a motor ship of the mass project 26-37 are presented.

Keywords: curves of the intensity of changes in stability characteristics, threshold values of changes in stability characteristics, hydrostatic characteristics of the ship's hull, ship landing, static stability diagram

АВТОМАТИЗАЦИЯ РАСЧЕТА ОПРЕДЕЛЕНИЯ КООРДИНАТ И ОТКЛОНЕНИЯ СРЕДНЕЙ ТОЧКИ ПОПАДАНИЯ ПО ФОТОИЗОБРАЖЕНИЮ

АО «ЦНИИ «Буревестник», г. Нижний Новгород, АО «Веда-проект» г. Москва

В статье рассматриваются основы разработки программного обеспечения для автоматизации процесса расчета баллистических параметров таких как: координаты средней точки попадания (СТП) и срединного отклонения (СО) по данным, полученным с фотоизображения.

Ключевые слова: системный подход, автоматизация расчетов, средняя точка попадания, срединное отклонение.

Проблема недостаточной автоматизации процессов получения и обработки баллистических данных считается одной из ключевых, влияющих на множество важных показателей при стрельбе.

В данной статье процесс автоматизации рассматривается с точки зрения использования системного подхода, т.е. выявление проблемы, анализ теоретического материала для ее решения, проработка вариантов решения, создание алгоритма работы и написание программного кода, проведение отладочных работ и выявление недостатков во время проведения отладочных работ.

После прицеливания в определенную точку оператор проводит серию выстрелов по мишени, вследствие этого на ней образуются пробоины (далее по тексту точки), занимающие некоторую площадь, которая называется площадью рассеивания. [1] (Рис. 1)

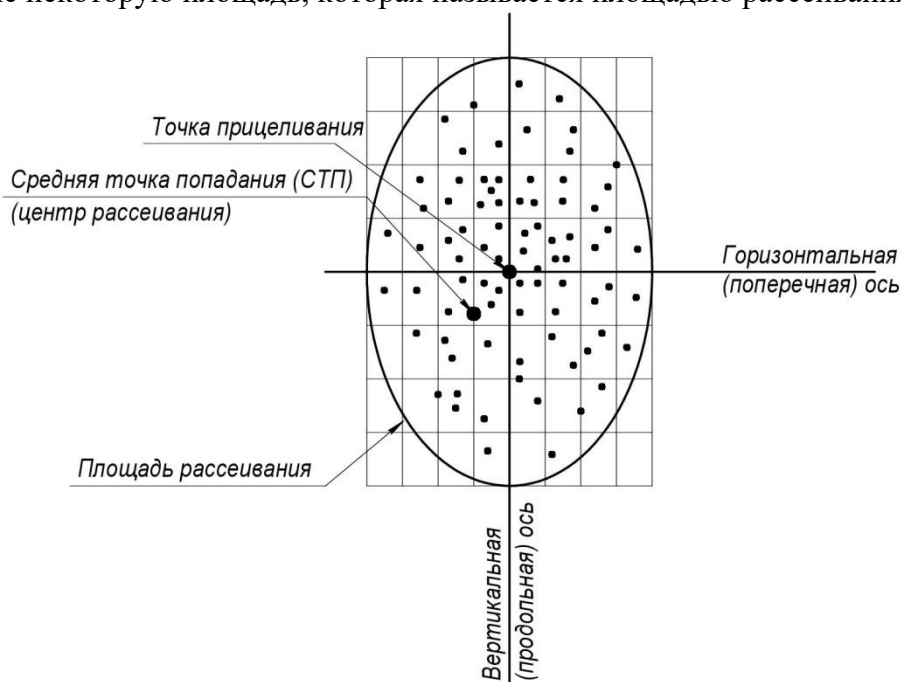


Рис. 1. Площадь рассеивания точек

Точка на данной площади, которая будет средней по отношению ко всем остальным, - называется средней точкой попадания (СТП). [1] Её координаты можно определить расчетным методом. Для этого необходимо провести через точку прицеливания вертикальную линию (продольная ось) и горизонтальную (поперечная ось).

Измерить кратчайшее расстояние от каждой пробойны до точки прицеливания, и сложив данные значения по каждой из осей, разделить каждую из сумм на число пробойн. Полученные значения определяют удаление СТП от указанных линий. [2]

Алгебраически это можно записать как:

$$X_{СТП} = \frac{\sum_{i=0}^n X_i}{n}, \quad (1)$$

где X - значения координат проекций на горизонтальную ось, n - количество точек. [3]

Не менее важной характеристикой является определение СО т.е. отклонения координат каждой точки по отношению к СТП.

Иными словами:

$$SO = \frac{3}{\pi} \cdot 0,6745 \cdot \sqrt{\frac{\sum_{i=1}^n (X_i - X_{СТП})^2}{n-1}}, \quad (2)$$

где SO-срединное отклонение, Xn - значения исследуемой величины.

В практической деятельности недостатком определения вышеперечисленных параметров является высокая трудоемкость получения большого массива исходных данных и их последующая обработка. Суть проблемы заключается в том, что на данный момент не разработано полностью автоматического преобразования графической информации с бумажных носителей (мишени) – в электронную форму необходимую для использования в баллистических вычислениях.

Известны следующие варианты решения данной проблемы:

1. Ручной счет. Данный способ является наиболее трудоёмким и используется до сих пор, из-за отсутствия других альтернатив в условиях требующих получения результатов. Недостатками данного метода являются: большие затраты времени на получение и обработку данных, а также необходимость последующего повторного их переноса в электронный вид.

2. Полуавтоматический расчет. Данный способ сейчас наиболее часто используется т.к. позволяет частично автоматизировать работу оператора. Его суть состоит в том, чтобы полученное фотоизображение поместить в графическую программу и провести измерения расстояния от точки прицеливания до центров точек.

Для выполнения расчета СТП и СО оператор вручную переносит полученные данные в файл Excel с заранее составленным в нем алгоритмом расчета. Недостатками данного метода является: высокая трудоемкость при получении и переносе исходных данных.

Для устранения вышеперечисленных недостатков разработана «Программа расчета баллистических параметров по фотоизображению».

Алгоритм программы выполнен с применением одного из методов системного подхода - «Дерево целей» [4]. Его горизонтальная развертка представлена на рис. 2.

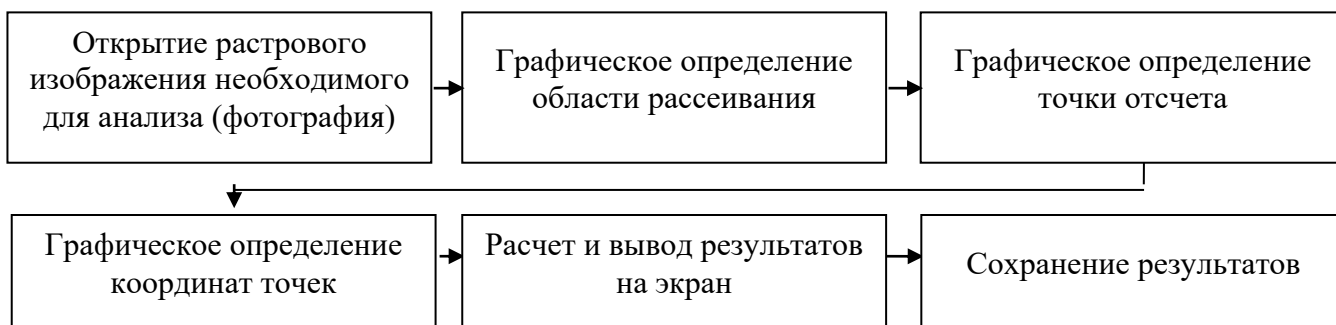


Рис. 2. Алгоритм работы программы

Описание работы программы

После запуска программы перед пользователем открывается диалоговое окно №1 (рис.3).



Рис. 3. Диалоговое окно №1

В открывшемся окне оператору необходимо указать значение дальности и выполнить загрузку необходимого фотоизображения, его пример показан на рисунке №4.

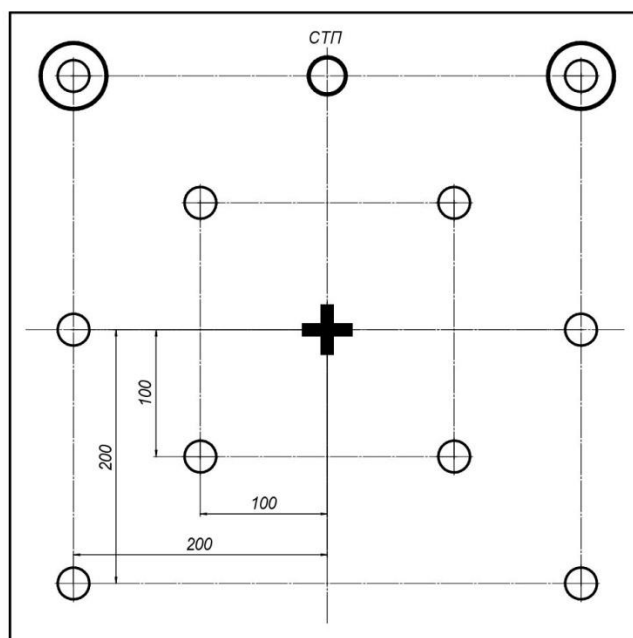


Рис. 4. Диалоговое окно №2

После этого определить точку отсчета и координаты пробойн (Рисунок 2). При этом, для удобства оператора, программой производится визуализация СТП с одновременным выводом информации координат СТП и СО в соответствующую область окна №1.

По окончании работ реализована возможность сохранения полученных результатов в файл Excel.

Выводы и дальнейшие перспективы исследования.

Практическое использование данного программного комплекса получило хорошую оценку пользователей, но и выявило ряд недостатков:

- отсутствует возможность использования фотоизображений с искаженной перспективой. Решение данной задачи возможно посредством разработки и внедрения математического алгоритма расчета обратного преобразования перспективы;

- недостаточно оперативное получение первичных данных посредством оператора. Использование технологии Convolutional Neural Network (CNN) смогло бы решить данную проблему.

Библиографический список

1. Наставление по стрелковому делу. Основы стрельбы из стрелкового оружия, Главное управление боевой подготовки, Астана, 2002 г., стр. 3-78.
2. Огневая подготовка: Учебник. - Москва. Воениздат, 2008 стр. 415
3. Венцель Е.С. Теория вероятности. Учебное издание. – Москва. Издательство «Высшая школа», 2006 г. (стр.240 10.3.25)
4. Волгушева А.А. «Дерево целей» - Москва, 2009. [Электронный ресурс]. – Режим доступа:
https://studme.org/207995/matematika_himiya_fizik/chislovy_e_harakteristiki_rasseyaniya_sluchaynoy_velichiny

E. Siatoikin, D. Lekomtsev

AUTOMATION OF CALCULATION WHEN DETERMINING COORDINATES AND DEVIATION OF THE MEAN POINT OF IMPACT BY PHOTO IMAGE

*BUREVESTNIK Central Research Institute JSC, Nizhny Novgorod;
VEDA-PROJECT JSC, Moscow.*

The article discusses the basics of software development for automatic calculation of ballistic parameters, such as coordinates of the mean point of impact (MPI) and probable circular error (CEP), based on the data taken from a photographic image.

Keywords: systems approach, automation of calculations, mean point of impact, probable circular error.

АВТОМАТИЗИРОВАННОЕ ПРЕОБРАЗОВАНИЕ В МОДЕЛЬ ЗНАНИЙ ОСД

Нижегородский государственный технический университет им. Р.Е. Алексеева

С появлением мощных вычислительных аппаратных комплексов появилась возможность создавать продвинутые искусственные интеллекты, которые существенно упрощают жизнь человека. Одной из задач искусственного интеллекта является обработка и анализ текста. В статье рассматривается исследование автоматического преобразования предложений в модель знаний ОСД.

Ключевые слова: анализ текстов, искусственный интеллект, модели знаний.

Введение

Технологический прогресс не стоит на месте. С появлением мощных вычислительных аппаратных комплексов появилась возможность создавать продвинутые искусственные интеллекты. Использование нейросетей дало возможность создавать системы способные отвечать на вопросы [8]. Однако данные системы не способны давать точные ответ на вопрос при условии отсутствия точной информации в базе знаний. При отсутствии точной информации они «обобщают» знания, дают ответ, но данная информация вряд ли может считаться ответом с точки зрения человеческого логического восприятия [8]. Данную проблему, возможно, решить с помощью ОСД сетей.

ОСД сеть

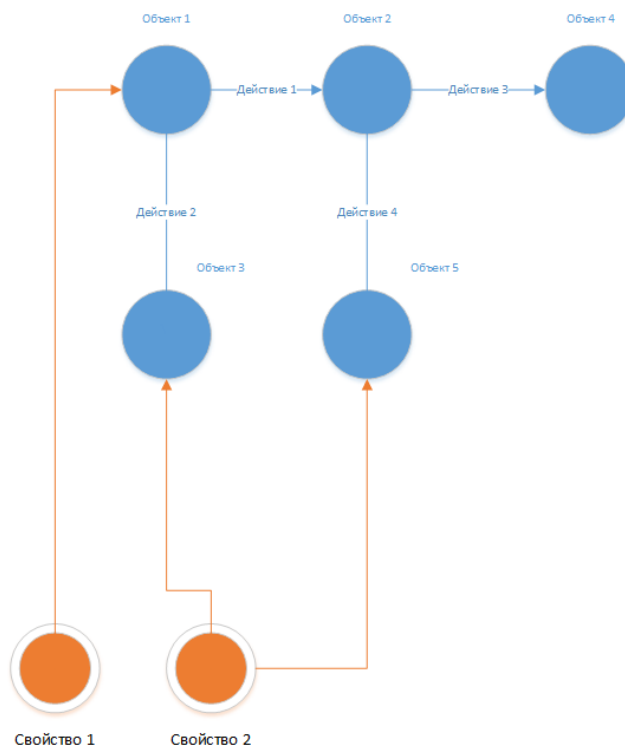


Рис. 1. Модель ОСД сети

ОСД аббревиатура означает: О-объект, С-свойство, Д-действие. ОСД – модель знаний, описывает взаимодействие объектов в тексте.

За объект модели принимается - имя существительное, за свойство - имя прилагательное, за действие – глагол.

Данную модель знаний можно представить в ориентированном и неориентированном графе. Вершина графа – объект, рёбра или дуги- действие. Свойство - это дополнительный элемент системы, который может быть связан с множеством объектов (рис. 1).

Авторы статьи выдвигают гипотезу о том, что информация в предложении разделяется на два типа:

1. Системная информация. Системной информацией являются объекты и их взаимосвязи в модели ОСД объект - действие - объект. В русском языке существительное либо местоимение, глагол, существительное либо местоимение.
2. Предметная информация. Предметной информацией являются свойства в модели ОСД -свойство. В русском языке это прилагательные, прилагательные, деепричастные и причастные обороты.

Для преобразования в модель знаний ОСД необходимо определить часть речи слова, в зависимости от этого слово ассоциируется ролью в модели ОСД.

Таблица 1

Часть речи	Роль в модели ОСД
<u>Имя существительное</u>	объект
<u>Имя числительное</u>	исключён
<u>Имя прилагательное</u>	свойство
<u>Местоимение</u>	метаданные
<u>Глагол</u>	действие
<u>Наречие</u>	свойство
<u>Предикатив</u>	свойство
<u>Причастие</u>	свойство
<u>Деепричастие</u>	свойство
<u>Предлог</u>	исключён
<u>Союз</u> («и» , «или», также символ «,») при условии формулы: [модельный тип] [,] [модельный тип]	логические метаданные
<u>Частица</u>	свойство
<u>Модальные слова</u>	исключён
<u>Междометия</u>	исключён

Метаданные позволяют создать связь с объектом из предыдущего предложения.

В модели ОСД действует правило, если в предложении последовательно встречаются два существительных, то они преобразуются в один объект ОСД.

Пример предложения из статьи о мозге человека

Мозг определяется как физическая и биологическая материя, содержащаяся в пределах черепа и ответственная за основные электрохимические и биоэлектрические нейронные процессы.

Первой задачей, является в определении системной информации, главного триплета «объект – действие –объект». Данная информация описывает взаимодействие между объектами. Именно в таких словах содержится функционально полезная информация для понимания текста.

1. слово: мозг; часть речи: существительное; роль в модели ОСД: объект;
2. слово: определяется; часть речи: глагол; роль в модели ОСД: действие;

3. слово: материя; часть речи: существительное; роль в модели ОСД: объект.

Пример разбора по модели ОСД, системной информации:

[мозг]-определяется-[материя].

Второй задачей, является определение предметной информации, вторичных цепочек взаимодействия. Для такого типа информации характерны цепочки связи:

1. объект – свойство;
2. объект – свойство – объект.

В модели ОСД к свойствам относятся такие типы речи: прилагательное, числительное, наречие, предикатив, причастие, деепричастие.

Пример разбора по модели ОСД, предметной информации

{ физическая }-[материя];

{ биологическая }-[материя];

[материя]-{ содержащаяся }-[пределах черепа];

[материя]-{ ответственная }-[процессы];

[материя]-{ основные }-[процессы];

[материя]-{ электрохимические }-[процессы];

[материя]-{ биоэлектрические }-[процессы];

[материя]-{ нейронные }-[процессы].

К основной проблеме классификации в модели ОСД относится множественное определение частей речи относительно символической последовательности слова.

Пример:

1. морю, как объект модели ОСД: вода идёт к морю;
2. морю, как действие модели ОСД: я морю тараканов.

Решением данной проблемы служит: определение возможных вариантов, а затем решение задачи прогнозирования. Под задачей прогнозирования понимается определение соседнего слова, и переопределение неоднозначного слова по формуле:

[объект]-действие-[объект].

Исследование результатов работы

В результате исследования была разработана программа, способная автоматически классифицировать слова по модели ОСД. Данная программа использует онлайн словари сервисы: викисловарь, рустекст, а также обладает алгоритмами определения части речи по окончанию слова.

Примеры автоматической классификации 1:

В результате исследования была разработана программа, способная автоматически классифицировать слова по модели.

Результат анализа:

0.1 - результате - (результат) Object | Status data =Reliable;

0.2 - исследования - (исследование) Object | Status data =Reliable;

0.0 - В - (в) Excluded | Status data =Reliable;

0.5 - программа - (программа) Object | Status data =Reliable;

0.6 - , - (,) LogicMetadata | Status data =Reliable;

0.3 - была - (быть) Action | Status data =Reliable;

0.4 - разработана - (разработать) Property | Status data =Reliable;

0.7 - способная - (способный) Property | Status data =Reliable;

0.9 - классифицировать - (классифицировать) Action | Status data =Reliable;

0.8 - автоматически - (автоматически) Property | Status data =Reliable;

0.11 - по - (по) Excluded | Status data =Reliable;

0.10 - слова - (слова) Object | Status data =Reliable;

0.12 - модели - (модель) Object | Status data =Reliable.

Примеры автоматической классификации 2:

Данная программа использует онлайн словари сервисы: викисловарь, рустекст, а также обладает алгоритмами определения части речи по окончанию слова.

Результат анализа:

- 0.1 - программа - (программа) Object | Status data =Reliable;
- 0.2 - использует - (использовать) Action | Status data =Reliable;
- 0.0 - Данная - (данный) Property | Status data =Reliable;
- 0.5 - сервисы - (сервис) Object | Status data =Reliable;
- 0.6 - викисловарь - (викисловарь) Object | Status data =Reliable;
- 0.3 - онлайн - (онлайн) Object | Status data =Reliable;
- 0.4 - словари - (словарь) Object | Status data =Reliable;
- 0.11 - также - (также) Excluded | Status data =Reliable;
- 0.8 - рустекст - (рустекст) Object | Status data =Reliable;
- 0.7 - , - () LogicMetadata | Status data =Reliable;
- 0.10 - а - (а) Undefined | Status data =Reliable;
- 0.9 - , - () LogicMetadata | Status data =Reliable;
- 0.12 - обладает - (обладать) Action | Status data =Reliable;
- 0.10 - а - (а) Excluded | Status data =Reliable;
- 0.13 - алгоритмами - (алгоритм) Object | Status data =Reliable;
- 0.14 - определения - (определение) Object | Status data =Reliable;
- 0.17 - по - (по) Excluded | Status data =Reliable;
- 0.16 - речи - (речь) Object | Status data =Reliable;
- 0.15 - части - (часть) Action | Status data =Reliable;
- 0.19 - слова - (слова) Object | Status data =Reliable;
- 0.18 - окончанию - (окончание) Object | Status data =Reliable.

Автоматизированный алгоритм классификации формирует ряд гипотез определения слов. После запускается алгоритм переопределения логических метаданных. Если логические метаданные не удовлетворяют условию, соседей с одинаковым типом модели ОСД, то слово исключается из обработки.

Заключение

На данный момент разработана программа, описанная выше. Есть методы и способы программного улучшения алгоритмов определения модели ОСД. В случае неоднозначной классификации слова, можно запросить пользователя об уточнении данных.

Возможность создания интеллектуальной экспертной системы с возможностью автоматического расширения сетевой структуры на базе модели ОСД представляется реальным, так как электронные словари имеют большую базу слов, альтернативные вариации слов и возможность расширения базы слов с помощью пользователей.

Проблема неоднозначности слова гипотетически высоковероятно решается с помощью прогнозирования и баз словосочетаний.

Модель ОСД имеет ряд преимуществ, одним из которых является получение системной информации. Взаимосвязи объект, действие, объект подпадает под классическое определение системы, что в свою очередь даёт возможность свободного интерпретирования данных из сети ОСД.

Библиографический список

1. Бронфельд Г.Б. Инженерные основы моделей знаний: учеб. Н. Новгород: НГТУ им. Р.Е. Алексеева, 2016. 139 с.
2. Бронфельд Г.Б. Основы искусственного интеллекта: учеб. пособие. Н. Новгород: НГТУ им. Р.Е. Алексеева, 2014. – 253 с.
3. Ландэ Д. В. Основы интеграции информационных потоков / Ландэ Д. В.//Информационные потоки. - 2006. - 240 с.

4. Комарцова Л.Г., Максимов А.В. Нейрокомпьютеры: учебное пособие для вузов. – 2-е изд., перераб. И доп. – М.: Изд-во МГТУ им. Н.Э. Баумана, 2004. – С. 210-213.
5. Хайкин С. Нейронные сети: полный курс, 2-е издание, исправленное.: Пер. с англ. – М.: ООО «И.Д. Вильямс», 2006. – 1104 с.: ил. – Парал. тит. англ. – С. 6–22.
6. Осовский С. Нейронные сети для обработки информации / Пер. с польского И.Д. Рудинского. – М.: Финансы и статистика, 2002. – 344 с.
7. Тарков М.С. Нейрокомпьютерные системы: Учебное пособие / М.С. Тарков. – М.: Интернет-Университет Информационных Технологий; БИНОМ. Лаборатория знаний, 2006. – 142 с.
8. Круглов В.В., Борисов В.В. Искусственные нейронные сети. Теория и практика. – 2-е изд., стереотип. – М.: Горячая линия-Телеком, 2002–322 с.

A.S. Makarov, V.P. Khranilov

AUTOMATED CONVERSION TO AN OPA KNOWLEDGE MODEL

Nizhny Novgorod state technical university n. a. R.E. Alekseev

Purpose: This article is dedicated to a program that tests the OPA knowledge model.

Design/methodology/approach: The design of the program closely follows object-oriented programming paradigm, exposing user friendly interface to user and extensible API to developer.

Findings: The program analyzes the text and makes a logical conclusion.

Research limitations/implications: The program architecture is flexible enough to provide engineers with a convenient way to create new features in order to extend system functionality.

Originality/value: The originality of the program is determined using the OPA model.

Keywords: text analysis, artificial intelligence, knowledge models.

**РАЗРАБОТКА СИСТЕМЫ ПОДДЕРЖКИ ПРОЦЕССА ВЫБОРА КЛИЕНТОМ
АВТОМОБИЛЯ**

*ООО Харман, г. Н. Новгород¹
НГТУ им. Р.Е. Алексеева, г. Н. Новгород²*

В данной работе рассматривается разработка вспомогательной системы, нацеленной на информационную поддержку пользователя при выборе автомобиля. Рассматриваются возможный функционал системы и её архитектурный тип.

Ключевые слова: многокритериальная оценка, экспертные системы, автомобили, web-приложение.

Актуальность: автомобильный рынок в настоящее время очень развит и может предложить потребителю автомобиль почти любой комплектации и ценовой категории. Но такое изобилие создаёт проблему выбора автомобиля - существуют клиенты, которым нужен автомобиль определённой марки с определёнными характеристиками, т.н. «автомобиль мечты», но затем покупают не то, что хотели по разным причинам: пришли в автосалон и не хватило денежных средств или их уговорили купить то, что в данный момент «на бонусе» для менеджера и прочее. А другие, не имея опыта покупки, вождения или имеющих мало знаний о технических характеристиках автомобиля, теряются в столь огромном выборе, не зная, что взять или берут не самый подходящий и не самый качественный вариант, поддаваясь уговорам работников автосалонов. И те, и другие в итоге тратят много времени на поиск, консультацию и итоговый выбор автомобиля. Разработка вспомогательной системы нацелена, на основе опроса, на генерацию списка наиболее подходящих автомобилей потребителю «не выходя из дома». Это экономит его время и нервы.

Характеристики процесса выбора: процесс работы представляет собой ввод пользователем в окне программы веса, от 0 до 100 следующих параметров автомобиля:

- безопасность;
- цена;
- комфорт;
- интерьер;
- экстерьер;
- гаджеты и устройства для обеспечения комфорта.

В дальнейшем число параметров может меняться в большую или меньшую сторону, в зависимости от решений дальнейшего развития архитектуры.

Параметр «безопасность» представляет собой эффективность систем автомобиля в области предотвращения, уменьшения вреда ДТП.

В параметр цены входит только цена автомобиля, без дополнительных затрат в виде запчастей, так как вероятность поломки любой детали автомобиля случайна, а цена на запчасти тоже имеет частое свойство колебаться.

Последние три параметра представляют собой «подпараметры» комфорта, так как они схожи по значению и для отдельных аргументов их роль в автомобиле недостаточна.

Опрос пользователя по таким параметрам удобен и не требует каких-либо дополнительных технических знаний с его стороны. Метод многокритериальной оценки используется по той причине, что для корректного анализа всех вариантов автомобилей использование одного какого-либо параметра недостаточно и всегда будет требоваться их набор.

После ввода весов каждого критерия для каждого автомобиля рассчитывается многокритериальная оценка, используя весовые коэффициенты. Формула следующая.

$$F=(S*WS+Y*WY+I*WI + Int*WInt + Ext*WExt)/100,$$

где S – значение критерия стоимости; Y – значение критерия комфорта; I – значение критерия безопасности; Int - значение критерия интерьера; Ext - значение критерия экстерьера; WS – вес критерия стоимости; WY – вес критерия эффективности; WI – вес критерия охвата; WInt - вес критерия интерьера; WExt - вес критерия экстерьера.

Далее значения сортируются и отбираются несколько автомобилей с самым большим числовым значением оценки.

Для более точного определения моделей проводится небольшой опрос предпочтений, например: наличие кондиционера/печки; предпочитаемый цвет; тип двигателя и прочее.

Исходя из ответов, начальный список отобранных автомобилей может меняться по более точным предпочтениям пользователя, выдвигая на выбор не вошедшие на этапе вычисления многокритериальной оценки автомобили.

Архитектуры системы

Наиболее подходящим вариантом архитектуры является разработка в виде экспертной системы (ЭС). Проводя поиск решения, ЭС использует заранее созданную базу знаний, которая может корректироваться и обновляться по мере необходимости.

Систему можно разбить условно на 3 части - интерактивная, логическая и хранилище данных.

Интерактивная часть представляет собой пользовательский интерфейс, в котором пользователь вводит необходимую для обработки запроса информацию - вес критериев, а также ответы на небольшой опрос. Также окно выдаёт список автомобилей, наиболее подходящих по введённым пользователем весам критериев.

Логическая часть представляет собой сервер, отвечающий за вычисление значений критериев каждого автомобиля (если требуется) и высчитывание многокритериальной оценки по формуле выше, с учётом проведённого опроса. Затем данные сортируются и передаются интерактивной части.

Данные представляет собой хранилище, БД с информацией о каждом автомобиле. Каждый автомобиль будет содержать следующую информацию: название, модель; список установленных систем; значение критериев - цена, числовые характеристики систем и пр.

Такая архитектура системы позволяет реализовать многопоточность, где каждый поток будет обрабатывать свою часть данных из БД и выводить результат в общий логический блок для сортировки всех обработанных данных. Схематично архитектуру можно представить рисунком ниже:

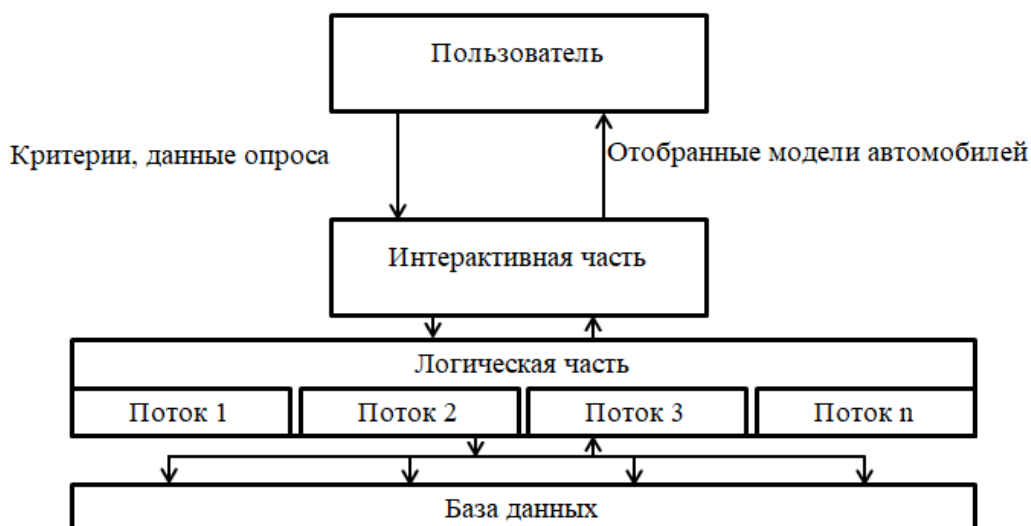


Рис. 1. «Архитектуры системы»

Такую систему можно разработать в нескольких видах:

- 1) Приложение на компьютер (ОС Windows/Linux);
- 2) Самостоятельное приложение на телефон (ОС Android/iOS);
- 3) Веб-приложение (нет привязки к ОС).

Первый вид не подходит, так как в современных реалиях пользователи перешли на работу с веб-приложениями, или же устанавливают нужное приложение себе на телефон.

Разработка приложения на телефон зависит от того, под какую ОС будет вестись разработка - от этого зависит выбор программных средств. Так для ОС Android разработка приложения будет на ЯП Java, а используемое IDE - Android Studio. Также необходим телефон с актуальной версией ОС для тестов. Для iOS - ЯП Swift или Objective-C, используемое IDE - Xcode, необходим один из продуктов Apple - MAC, iPhone для тестов.

Второй путь сложнее из-за дороговизны тестовых устройств, а саму iOS необходимо детально изучать на разрешённые команды и операции в пользовательских приложениях.

Веб-приложение — состоит из HTML страниц. Но главным отличием от обычных сайтов является то, что HTML-страницы генерируются на лету в зависимости от запроса пользователя. Почтовые клиенты, соцсети, поисковики, интернет-магазины, онлайн-программы для бизнеса, это все веб-приложения.

Веб-приложения можно разделить на несколько типов, в зависимости от разных сочетаний его основных составляющих:

Backend (бэкенд или серверная часть приложения) работает на удаленном компьютере, который может находиться где угодно. Она может быть написана на разных языках программирования: PHP, Python, Ruby, C# и других. Если создавать приложение используя только серверную часть, то в результате любых переходов между разделами, отправки форм, обновления данных, сервером будет генерироваться новый HTML-файл и страница в браузере будет перезагружаться.

Frontend (фронтенд или клиентская часть приложения) выполняется в браузере пользователя. Эта часть написана на языке программирования Javascript. Приложение может состоять только из клиентской части, если не требуется хранить данные пользователя дольше одной сессии. Это могут быть, например, фоторедакторы или простые игрушки.

Single page application (SPA или одностраничное приложение). Более интересный вариант, когда используются и бэкенд и фронтенд. С помощью их взаимодействия можно создать приложение, которое будет работать совсем без перезагрузок страницы в браузере. Или в упрощенном варианте, когда переходы между разделами вызывают перезагрузки, но любые действия в разделе обходятся без них.

Для работы системы необходимо разработать backend и frontend часть, так как системе нужно хранить и обрабатывать данные на стороне сервера, а также отображать их в реальном времени без перезагрузки на стороне клиента.

Было принято решение разработать систему в виде веб-приложения, так как сейчас это очень модное направление, и оно удобно для пользователя, так как можно получить информацию, находясь за домашним компьютером, и за телефоном, при этом не устанавливая что-либо на вышеописанные устройства.

Разработку серверной части можно разделить на условные 3 группы языков и фреймворков:

— Ruby фреймворки (Ruby on Rails, Sinatra)

Ruby on Rails — фреймворк, написанный на языке программирования Ruby, реализует архитектурный шаблон Model-View-Controller для веб-приложений, а также обеспечивает их интеграцию с веб-сервером и сервером баз данных.

— Python (Django)

Python — высокоуровневый язык программирования общего назначения с динамической типизацией и автоматическим управлением памятью, ориентированный на

повышение производительности разработчика, читаемости кода и его качества, а также на обеспечение переносимости написанных на нём программ.

Django — свободный фреймворк для веб-приложений на языке Python, использующий шаблон проектирования MVC. В Django для работы с БД используются модели (model). Они позволяют описывать таблицы и делать запросы на привычном разработчику питоне, что гораздо удобнее. За это удобство приходится платить: такие запросы медленнее и ограничены в возможностях по сравнению с использованием чистого SQL.

— PHP (Yii, Symfony)

PHP — скриптовый язык общего назначения, интенсивно применяемый для разработки веб-приложений.

Symfony — свободный фреймворк, написанный на PHP. Symfony предлагает быструю разработку и управление веб-приложениями, позволяет легко решать рутинные задачи веб-программиста. Работает только с PHP 5 и выше. Имеет поддержку множества баз данных.

Yii — основанный на компонентной структуре PHP-фреймворк для быстрой разработки крупных веб-приложений. Он позволяет максимально применить концепцию повторного использования кода и может существенно ускорить процесс веб-разработки. Разработку клиентской части можно разделить на следующие группы фреймворков:

React — JavaScript-библиотека с открытым исходным кодом для разработки пользовательских интерфейсов. React разрабатывается и поддерживается Facebook, Instagram и сообществом отдельных разработчиков и корпораций. React может использоваться для разработки одностраничных и мобильных приложений.

Vue.js — JavaScript-фреймворк с открытым исходным кодом для создания пользовательских интерфейсов. Легко интегрируется в проекты с использованием других JavaScript-библиотек. Может функционировать как веб-фреймворк для разработки одностраничных приложений в реактивном стиле.

Angular — открытая и свободная платформа для разработки веб-приложений, написанная на языке TypeScript, разрабатываемая командой из компании Google, а также сообществом разработчиков из различных компаний. Angular — полностью переписанный фреймворк от той же команды, которая написала AngularJS.

Ember.js — свободный каркас веб-приложений на JavaScript, реализующий шаблон MVC и нацеленный на упрощение создания масштабируемых одностраничных веб-приложений.

Для разработки был выбран язык Python с фреймворком Django для серверной части и платформа Angular для клиентской части, так как язык Python более привычен и удобен в использовании, имея множество примеров использования. Также имеются библиотеки для работы с различными базами данных, а Angular имеет поддержку TypeScript, что более удобно, чем JavaScript.

Выбор СУБД для хранения данных:

MongoDB — документоориентированная система управления базами данных, не требующая описания схемы таблиц. Считается одним из классических примеров NoSQL-систем, использует JSON-подобные документы и схему базы данных.

MySQL — свободная реляционная система управления базами данных. Разработку и поддержку MySQL осуществляет корпорация Oracle, получившая права на торговую марку вместе с поглощённой Sun Microsystems, которая ранее приобрела шведскую компанию MySQL AB.

SQLite — компактная встраиваемая СУБД. Исходный код библиотеки передан в общественное достояние.

PostgreSQL — свободная объектно-реляционная система управления базами данных. Существует в реализациях для множества UNIX-подобных платформ, включая

AIX, различные BSD-системы, HP-UX, IRIX, Linux, macOS, Solaris/OpenSolaris, Tru64, QNX, а также для Microsoft Windows.

Была выбрана СУБД MySQL так как она имеет удобный клиент для работы с БД, а также свой собственный сервер. Также ЯП Python имеет библиотеки для работы с этим сервером.

Таким образом, разрабатываемое приложение будет в виде web-приложения и будет разделено на клиентскую и серверную часть, которая в свою очередь будет работать с БД на языке SQL. Клиентская часть будет разрабатываться на платформе Angular, серверная на Python с использованием фреймворка Django. База данных будет создана в СУБД MySQL, управление будет производиться серверной частью.

Также система будет иметь настраиваемую интеллектуальную поддержку для помощи в выборе автомобиля в виде ввода весов и прохождения опроса.

N.D. TsyganovA. I. Markovnina

DEVELOPMENT OF A SUPPORT SYSTEM FOR THE PROCESS OF CHOOSING A CAR BY A CLIENT

Nizhny Novgorod state technical university n.a. R.E. Alekseev

Purpose: The development of an auxiliary system is aimed, based on a survey, at generating a list of the most suitable vehicles for the users “from home”.

Design/methodology/approach: web application.

Findings: This document discusses the development of an auxiliary system aimed at providing information support to the user when choosing a car. The possible functionality of the system and its architectural type are considered.

Research limitations/implications: system can be used for many types of vehicle.

Originality/value: application can save user time and hassle for choosing a vehicle.

Keywords: multi-criteria assessment, expert systems, cars, web application.

А.О. Семенов, А.А. Штанюк

ПОСТРОЕНИЕ ГРАФОВОЙ МОДЕЛИ ТОПОЛОГИИ ИНТЕГРАЛЬНЫХ СХЕМ НА ОСНОВЕ ДАННЫХ ИЗ GDSII ФАЙЛА

Нижегородский государственный технический университет им. Р.Е. Алексеева

В статье рассматривается проблема анализа и верификации топологий ИС, хранящихся в виде файлов, на примере файлов формата GDSII. Помимо GDSII существуют и другие достаточно распространенные форматы для хранения топологий ИС, что влечет за собой дополнительные сложности при анализе и сравнении топологий различных форматов. Возникает необходимость привести файлы топологий ИС к единому виду, для чего в работе предлагается использовать графовые представления. В данной работе описывается алгоритм преобразования файла топологии ИС в графовую модель на примере файлового формата GDSII. Графовые модели не только позволяют удобно сравнивать топологии ИС различных форматов между собой, но и открывают дополнительные возможности для анализа топологий ИС с помощью алгоритмов теории графов.

Ключевые слова: GDSII, сверхбольшая интегральная схема, топология интегральных схем, верификация, алгоритм, графовые модели.

Введение

Файловый формат GDS (Graphic Database System) был разработан организацией Calma Company (США) в начале 70-х годов прошлого века и на многие годы стал стандартным цифровым форматом файла для обмена данными по интегральным схемам и их топологиям [1]. Вторая версия, GDSII, была представлена в 1978 году. В течение многих лет этот формат был единственным в своем роде, и многие другие производители приняли его в своих системах.

Несмотря на то, что для замены устаревшего GDSII разрабатываются новые форматы, он до сих пор одним из основных способов описания топологий интегральных схем, используемых для построения чипа.

Как было сказано ранее, GDSII — это двоичный формат файла, представляющий плоские геометрические фигуры, текстовые метки и другую информацию о макете интегральной схемы в иерархической форме. Данные файла GDSII можно использовать для реконструкции всей геометрии интегральных схем или ее части, которая будет использоваться при обмене макетами, передаче рисунков между различными инструментами или создании фотомасок (фотошаблонов).

Несмотря на ограниченный набор функций и низкую плотность данных, он стал общепринятым отраслевым форматом для передачи данных компоновки ИС между инструментами проектирования разных производителей, которые работали с собственными форматами данных.

Объекты, содержащиеся в файле GDSII, группируются путем присвоения им числовых атрибутов, включая номер слоя, тип данных или текстовый тип. Хотя эти атрибуты были разработаны так, чтобы соответствовать слоям материала, используемым при изготовлении интегральных схем, их значение быстро стало более абстрактным, чтобы отразить способ проектирования физической схемы.

Файлы GDSII — это набор ячеек, которые могут содержать полигоны или другие ссылки на ячейки [2]. Эти ячейки, называемые структурами на языке GDSII, имеют буквенно-цифровые имена длиной до 32 символов.

Библиотека этих структур содержится в файле, который состоит из **заголовка библиотеки, последовательности структур и окончания библиотеки. Каждая структура в последовательности состоит из заголовка структуры, последовательности элементов и окончания структуры.**

Текстовым вариантом бинарного файла GDSII является GDS-TXT. Существует семь видов элементов GDSII [2, 3]:

- 1) граница (BOUNDARY) определяет заполненный многоугольник;
- 2) путь (PATH) определяет проводник;
- 3) ссылка (SREF) на структуру вызывает все элементы, которые находились в этой структуре;
- 4) ссылка на массив структур (AREF) вызывает массив структур и все элементы внутри структур;
- 5) текст (TEXT) предназначен для документации;
- 6) узел (NODE) определяет электрический путь;
- 7) прямоугольник (BOX) размещает прямоугольную геометрию.

Основным недостатком является то, что формат не оптимизирован для уменьшения размера файла, в то время как проекты часто превышают размер в 20 гигабайт. Более подробно недостатки GDSII описаны в статье [4].

Данные GDSII имеют иерархическую структуру, которая позволяет описывать миллионы или миллиарды полигонов в ИС. Иерархия очень важна, потому что в дизайне ИС, как правило, повторно используются одни и те же геометрические фигуры. Благодаря этому можно создавать компактные файлы для очень больших конструкций [5, 6].

На данный момент существуют несколько форматов файлов, используемых для хранения топологии интегральных схем. Одним из наиболее популярных форматов является OASIS, который более подробно рассмотрен в статье [7]. Файлы формата OASIS имеют значительно меньший размер, чем файлы формата GDSII. Формат DEF также относится к форматам, хранящим и описывающим топологии интегральных схем. Из-за большого количества форматов с различными элементами и средствами проектирования возникает потребность привести топологии интегральных схем к единой форме представления. Это позволит проще и удобнее анализировать и сравнивать между собой топологии ИС, хранящиеся в различных форматах, не привязываясь к специфике конкретного формата.

Так как данные в GDSII представлены в иерархической форме, то одним из наиболее эффективных решений проблем, связанных со сравнением данных хранящихся в файлах различных форматов, является представление элементов и связей между ними в виде графовых моделей. Так, в работе [8] авторы находят решение проблемы анализа и сравнения файлов формата DEF и GDSII в представлении данных в виде графа и в дальнейшей проверке графов на изоморфность.

Следует отметить, что построение графовых моделей позволяет не только привести файлы разных форматов к единому виду, но и открывает дополнительные возможности для анализа и верификации файлов с помощью различных графовых алгоритмов. Таким образом, в данной работе описывается алгоритм, позволяющий преобразовывать данные топологий интегральных схем, хранящихся в формате GDSII, в граф.

Алгоритм построения графа на основе файла формата GDSII

Данный алгоритм позволяет выявить наличие связей между всеми элементами структуры и построить граф связей. Алгоритм можно разделить на две фазы. В первой фазе выявляются все объекты, которые являются коннекторами, то есть соединяют между собой другие объекты. В роли коннекторов могут выступать следующие объекты: объект PATH или любой объект специального слоя, элементы которого служат контактными площадками. Информацию о найденных элементах-коннекторах

необходимо сохранять в памяти программы. Информация должна включать в себя номер слоя элемента, координаты элемента, а также типы и имена родителей. В роли родителей элемента могут выступать структура, SREF, AREF.

В первой фазе для всех элементов выбранной структуры необходимо выполнить следующие действия:

- 1) Если элемент находится в изоляционном слое, то он не рассматривается, так как является изолятором.
- 2) Если элемент является BOX или BOUNDARY, то проверяем номер его слоя:
 - a) если слой является специальным слоем, то элемент является контактной площадкой, т. е. коннектором;
 - b) если слой не является слоем контактных площадок, то элемент не является коннектором.
- 3) Если элемент является PATH, то он является коннектором.
- 4) Если элемент является SREF, т. е. ссылкой на другую структуру, то необходимо выполнить шаги 1-4 для всех элементов данной структуры.
- 5) Если элемент является AREF, т. е. массивом ссылок на другую структуру, то для каждого элемента данного массива необходимо выполнить шаг 4.

Таким образом, в конце первой фазы алгоритма получен набор элементов-коннекторов с их родителями. Данная информация необходима для выполнения второй фазы алгоритма.

Целью второй фазы является построение графа связей для выбранной структуры. Для этого необходимо выявить все связи элементов-коннекторов с другими элементами. Элементами-коннекторами могут являться элементы типа PATH или любые элементы, находящиеся в слое контактных площадок. Для выявления связей коннекторов структуры, находящихся в слое контактных площадок, необходимо выполнить следующие действия для всех элементов структуры:

- 1) если элемент находится в слое изоляторов, то он не рассматривается;
- 2) если элемент является тем же самым элементом, то переходим к следующему элементу;
- 3) если элемент является BOX или BOUNDARY, или PATH, то выполняется проверка пересечения коннектора с данным элементом;
- 4) если элемент является SREF, то проверяется пересечение этого элемента с коннектором. Если выявлен факт пересечения элементов, то необходимо выполнить следующие действия:
 - a) если граф не построен для структуры, на которую ссылается SREF, то выполняется алгоритм построения графа для данной структуры (1 и 2 фазы текущего алгоритма);
 - b) если граф построен для структуры, на которую ссылается SREF, то необходимо проверить, находится ли коннектор в данной структуре (по информации о родителях коннектора). Если коннектор находится в структуре, то необходимо из графа получить информацию о том, с какими элементами он связан. Если коннектор не принадлежит данной структуре, то выполняется проверка на пересечение коннектора со всеми элементами данной структуры;
- 5) если элемент является AREF, то проверяется пересечение этого элемента с коннектором. Если выявлен факт пересечения элементов, то необходимо рассмотреть каждый элемент массива ссылок AREF по отдельности как элементы SREF и выполнить для них шаг 4.

Чтобы выявить все связи коннектора типа PATH, необходимо выполнить следующие действия для каждого элемента структуры:

- 1) если элемент находится в слое изоляторов, то он не рассматривается;

- 2) если элемент является тем же самым элементом, то переходим к следующему элементу;
- 3) если элемент является BOX или BOUNDARY, или PATH то необходимо проверить, находится ли коннектор в одном слое с данным элементом, если это так, то выполняется проверка пересечения коннектора с данным элементом;
- 4) если элемент является SREF, то проверяется пересечение для тех групп элементов в SREF, которые находятся в том же слое, что и коннектор. Если выявлен факт пересечения, то необходимо выполнить следующие действия:
 - а) если граф не построен для структуры, на которую ссылается SREF, то выполняется алгоритм построения графа для данной структуры (1 и 2 фазы текущего алгоритма);
 - б) если граф построен для структуры, на которую ссылается SREF, то необходимо проверить, находится ли коннектор в данной структуре (по информации о родителях коннектора). Если коннектор находится в структуре, то необходимо из графа получить информацию о том, с какими элементами он связан. Если коннектор не принадлежит данной структуре, то выполняется проверка на пересечение коннектора со всеми элементами данной структуры;
- 5) если элемент является AREF, то проверяется пересечение коннектора с теми группами элементов в AREF, которые находятся в том же слое, что и коннектор. Если выявлен факт пересечения, то необходимо рассмотреть каждый элемент массива ссылок AREF по отдельности как элементы SREF и выполнить для них шаг 4.

Для демонстрации того, как должен работать описанный алгоритм, можно рассмотреть элементарный пример, изображенный на рис. 1. Из рисунка видно, что существует структура В, которой принадлежат элементы Boundary3 и Path3. Также структуре В принадлежит ссылка SREF A на структуру А, которой принадлежат элементы Boundary1, Boundary2, Path1, Path2. С помощью алгоритма, описанного выше, можно получить граф связей структуры В, который визуальнo отображен на рисунке 2. Вершинами графа являются элементы типа BOUNDARY, ребрами графа являются элементы типа PATH. Таким образом, выявлены все связи элементов структуры В между собой.

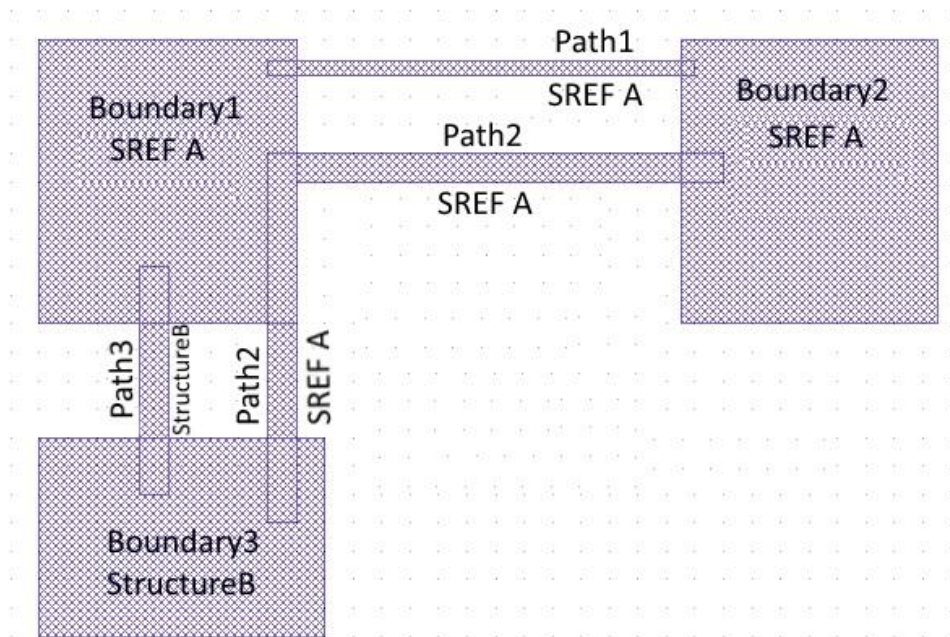


Рис. 1. Элементарный пример расположения элементов

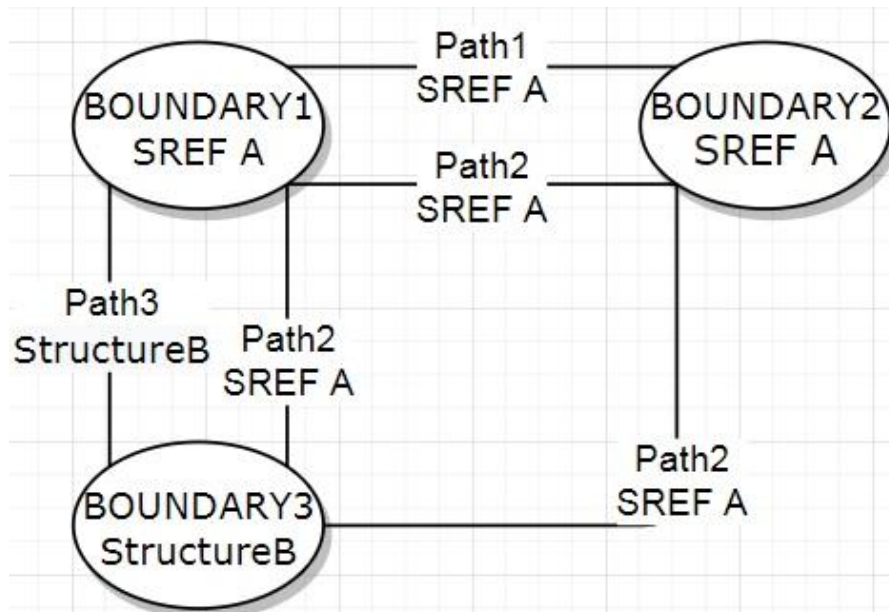


Рис. 2. Граф связей между элементами

Вывод

В данной статье был рассмотрен файловый формат GDSII и его особенности. Из-за большого количества форматов с различными элементами и средствами проектирования был предложен способ преобразования информации из файлов в графовые модели. Также приведен алгоритм преобразования данных из GDSII в граф, который упрощает анализ и проверку корректности топологий интегральных схем, а также сравнение топологий интегральных схем различных форматов. Анализ файлов, содержащих информацию о топологиях, приводится к анализу графов, что позволяет использовать различные алгоритмы теории графов и в то же время не требует от пользователя глубоких знаний архитектуры того или иного файлового формата.

Библиографический список

1. Bartolf H. Nanoscale-Precise Coordinate System: Scalable, GDSII-Design / H. Bartolf // Fluctuation Mechanisms in Superconductors. – Springer Spektrum, Wiesbaden, 2016 – с. 37-42.
2. Calma Corporation, GDS II Stream Format Manual (B97E060), Release 6.0 // Calma, February 1987. [Электронный ресурс]. – Режим доступа:
3. GDS_II_Stream_Format_Manual_6.0_Feb87.pdf (дата обращения: 26.10.2020).
4. GDSII [Электронный ресурс]. – Режим доступа: <https://sites.google.com/site/esdcat/gdsii> (дата обращения: 13.11.2020)
5. Штанюк А.А. Проблема анализа топологии интегральных схем на основе GDSII-файлов/ А.А. Штанюк, А.О. Семенов // Инженерные и информационные технологии, экономика и менеджмент в промышленности: сборник научных статей по итогам второй международной научной конференции. – Волгоград, 2020. с. 334 – 336.
6. Singla A. Verification of Physical Chip Layouts Using GDSII Design Data / A. Singla, B. Lippmann, H. Graeb // 2019 IEEE 4th International Verification and Security Workshop (IVSW). – IEEE, 2019. – с. 55-60.
7. Le C. Improved file sizes and cycle times through optimization of GDSII Stream / C. Le, D. Gariery // SPIE Proceedings, International Society for Optics and Photonics, 2005 – с. 351 – 358.

8. Штанюк А.А. Формат файлов OASIS для хранения топологии интегральных схем / А.А. Штанюк, А.О. Семенов // Приоритетные направления инновационной деятельности в промышленности: сборник научных статей по итогам двенадцатой международной научной конференции. – Казань, 2020. с. 106 – 107.
9. Галашов Д. А. Постановка задачи исследования алгоритмов проверки изоморфизма гиперграфов при анализе топологии интегральных схем на основе GDSII и DEF-файлов / Д. А. Галашов, А. А. Штанюк // Информационные системы и технологии ИСТ-2020: сборник трудов Международной научно-технической конференции. – Нижний Новгород, 2020. с. 743-749.

A.O. Semenov, A.A. Shtanyuk

BUILDING A TOPOLOGY GRAPH MODEL OF INTEGRATED CIRCUITS BASED ON DATA FROM A GDSII FILE

Nizhniy Novgorod state technical university n.a. R.E. Alekseev

The paper deals with the problem of analysis and verification of IC topologies stored in the GDSII format files. Besides GDSII there are other quite common formats for storing IC topologies, which entails additional difficulties in analyzing and comparing the topologies of different formats. There is a need to bring the IC topology files to a unified form, for which this paper proposes to use graphical representations. In this work the algorithm of conversion of IC topology file into graphical model is described on the example of GDSII file format. Graph models not only allow easy comparison of IC topologies in different formats, but also open additional possibilities for analysis of IC topologies by means of graph theory algorithms.

Keywords: GDSII, integrated circuit, integrated circuit topology, verification, algorithm, graph models.

ПРИНЦИПЫ ПОДДЕРЖКИ КОНВЕРТАЦИИ ДАННЫХ В РАЗНОРОДНЫХ АВТОМАТИЗИРОВАННЫХ СИСТЕМАХ С ПРИМЕНЕНИЕМ СРАВНЕНИЯ БЛИЗОСТИ СТРУКТУР ГЕОМЕТРИЧЕСКИХ МОДЕЛЕЙ

Нижегородский государственный технический университет им. Р.Е. Алексеева

Предложена методика сравнения близости структур геометрических моделей в разнородных автоматизированных системах с использованием графов, в рамках поддержки процесса конвертации между системами на примере CAD- и PDM-систем. От корректного взаимодействия разнородных автоматизированных систем, результативности руководства работами на стадии проектирования, введения в общую структуру единого информационного пространства зависит конкурентоспособность производственных предприятий.

Ключевые слова: геометрическая модель, формат передачи данных, конвертация параметров, CAD-система, PDM-система.

Введение

В рамках постоянного увеличения сложности конструкторских, инженерно-технических проектов, а также планов по планированию и реализации новой продукции предприятий, имеют неоспоримое преимущество такие промышленные предприятия, на которых первичными и хорошо отлаженными являются процессы планирования, разработки, проектирования, поставки, а также, непосредственно производства и сопровождения изделия. На рынке при необходимости появления нового продукта, промышленное предприятие должно за ограниченный срок создать требуемый рынком продукт и поставить его.

В процессе передачи данных в рамках крупных холдингов, между предприятиями при реализации технологий поддержки жизненного цикла изделий наблюдается переход к концепции единого информационного пространства (ЕИП) [1]. Переход к ЕИП является важной задачей для организаций и предприятий, что позволяет более быстро и эффективно управлять деятельностью сотрудников и принимать решения. Достаточно часто бывают распространенными состояния, когда часть учреждений, организаций автоматизированы только частично, на некоторых стадиях жизненного цикла. В таких случаях часто применяются самые разные программные комплексы. Могут применяться разные вендоры на стадиях жизненного цикла, часто на предприятиях бывает несколько поставщиков программного обеспечения в качестве разнородных автоматизированных систем. Форматы программных комплексов на предприятиях и различных крупных холдингах часто не совместимы, возможны дополнительные затраты при реализации информационной поддержки жизненного цикла (рис.1). В состав разнородных систем автоматизации входят системы автоматизированного проектирования (CAD-системы), системы управления данными об изделии (PDM-системы) [2]. Как показывает практика, на отечественных предприятиях применяется несколько CAD-систем под управлением PDM-систем, переход к единой автоматизированной системе затруднен в связи с большим объемом конструкторских данных и специализированных особенностей данных автоматизированных систем. От качества интеграции в общую структуру единого информационного пространства зависит конкурентоспособность предприятий [3].

Существуют проблемы при использовании программных пакетов, средств, комплексов различных вендоров, имеющих свои форматы при сохранении геометрических моделей (ГМ). Такие форматы могут быть закрытыми. Потери при передаче данных между автоматизированными системами с несовместимыми форматами

может привести к затратам на восстановление, важно это учитывать при передаче параметров между крупными холдингами.



Рис. 1. Информационная поддержка этапов жизненного цикла изделий

Методика сравнения близости структур ГМ

Каждое изделие имеет древовидную структуру, а также алгоритмы функционирования изделий предполагают проверку структуры, поэтому для сравнения близости графов требуется описать данную методику с применением графов [4]. Далее рассмотрены ГМ, исследование с которыми проведено. Первая – исходная ГМ, это некоторое изделие, для которого необходимо провести конвертацию. Вторая – результат конвертации в другую систему автоматизации, что получено на выходе некоторой операции с данными. Необходимо исследовать возможные проблемы при конвертации ГМ между системами автоматизации с применением теории графов и математически осуществить представление в виде графа типа «дерево». Разработана структура ГМ с применением графов, представлена в виде списка отдельных элементов [5], которые включили в себя параметры интеграции и параметры файла разнородных автоматизированных систем, а также геометрическую и атрибутивную информацию. Данная структура является источником данных для определения со структурных весовых коэффициентов элементов ГМ. Структура включает данные по интеграции систем автоматизации, а также дополнительные элементы, связанные с атрибутивными параметрами ГМ, что дополнительно реализовано в виде фреймов [6]. ГМ представляется структурой, включающей в себя дерево изделия, не содержащего циклов (N-арного дерева) и фреймы, содержащие параметры ГМ. Иерархическая структура изделия содержащую версии составных частей и характеристики этих частей, обозначается через граф $G = (X, A)$.

На начальном этапе сравнивается близость только деревьев ГМ на основании математического представления деревьев в виде матриц смежности. Матрица смежности представляет собой квадратную матрицу с логическими значениями (0 или 1). Информация о параметрах ГМ отражается в наличии рёбер графа, также при описании графов типа дерево для структуры ГМ, информация содержится в первую очередь в вершинах. Если пропадет вершина дерева, то пропадет и ребро. Для наглядности описываем часть сборки изначальной ГМ, представленной на рис. 2.

Матрица смежности представляет собой квадратную матрица размерностью $n \times n$, (где n – число вершин графа), однозначно представляющая его структуру, одна из вариаций построений графа в виде матрицы.

$A = \{a_{ij}\}$, $i, j = 1, 2, \dots, n$, а каждый элемент матрицы определяется следующим образом:
 $a_{ij} = 1$, если \exists дуга (x_i, x_j) ; $a_{ij} = 0$, если нет дуги (x_i, x_j) .

Матрицы графов сравниваются для того чтобы выявить визуально незаметное сходство структуры графов, в нашем случае, это сходство структуры ГМ до проведения процесса конвертации и после. Для сравнения графов используют их матричное представление. Из одной матрицы вычитается другая. Чем меньше полученный результат,

тем более похожи графы. Матрица смежности данной ГМ и бинарные значения матрицы представлены далее

$$A = \begin{pmatrix} a_{11} & a_{12} & a_{13} & a_{14} & a_{15} & a_{16} & a_{17} & a_{18} & a_{19} \\ a_{21} & a_{22} & a_{23} & a_{24} & a_{25} & a_{26} & a_{27} & a_{28} & a_{29} \\ a_{31} & a_{32} & a_{33} & a_{34} & a_{35} & a_{36} & a_{37} & a_{38} & a_{39} \\ a_{41} & a_{42} & a_{43} & a_{44} & a_{45} & a_{46} & a_{47} & a_{48} & a_{49} \\ a_{51} & a_{52} & a_{53} & a_{54} & a_{55} & a_{56} & a_{57} & a_{58} & a_{59} \\ a_{61} & a_{62} & a_{63} & a_{64} & a_{65} & a_{66} & a_{67} & a_{68} & a_{69} \\ a_{71} & a_{72} & a_{73} & a_{74} & a_{75} & a_{76} & a_{77} & a_{78} & a_{79} \\ a_{81} & a_{82} & a_{83} & a_{84} & a_{85} & a_{86} & a_{87} & a_{88} & a_{89} \\ a_{91} & a_{92} & a_{93} & a_{94} & a_{95} & a_{96} & a_{97} & a_{98} & a_{99} \end{pmatrix}, \quad A = \begin{pmatrix} 011000000 \\ 000100000 \\ 000011000 \\ 000000000 \\ 000000110 \\ 000000001 \\ 000000000 \\ 000000000 \\ 000000000 \\ 000000000 \end{pmatrix} \quad (1)$$

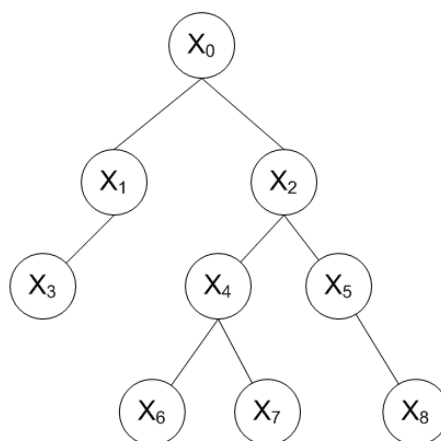


Рис. 2. Пример исходной ГМ

Матрица смежности представляет собой квадратную матрица размерностью $n \times n$, (где n – число вершин графа), однозначно представляющая его структуру, одна из вариаций построений графа в виде матрицы.

$A = \{a_{ij}\}$, $i, j = 1, 2, \dots, n$, а каждый элемент матрицы определяется следующим образом: $a_{ij} = 1$, если \exists дуга (x_i, x_j) ; $a_{ij} = 0$, если нет дуги (x_i, x_j) .

Матрицы графов сравниваются для того чтобы выявить визуально незаметное сходство структуры графов, в нашем случае, это сходство структуры ГМ до проведения процесса конвертации и после. Для сравнения графов используют их матричное представление. Из одной матрицы вычитается другая. Чем меньше полученный результат, тем более похожи графы. Матрица смежности данной ГМ и бинарные значения матрицы представлены далее

$$A = \begin{pmatrix} a_{11} & a_{12} & a_{13} & a_{14} & a_{15} & a_{16} & a_{17} & a_{18} & a_{19} \\ a_{21} & a_{22} & a_{23} & a_{24} & a_{25} & a_{26} & a_{27} & a_{28} & a_{29} \\ a_{31} & a_{32} & a_{33} & a_{34} & a_{35} & a_{36} & a_{37} & a_{38} & a_{39} \\ a_{41} & a_{42} & a_{43} & a_{44} & a_{45} & a_{46} & a_{47} & a_{48} & a_{49} \\ a_{51} & a_{52} & a_{53} & a_{54} & a_{55} & a_{56} & a_{57} & a_{58} & a_{59} \\ a_{61} & a_{62} & a_{63} & a_{64} & a_{65} & a_{66} & a_{67} & a_{68} & a_{69} \\ a_{71} & a_{72} & a_{73} & a_{74} & a_{75} & a_{76} & a_{77} & a_{78} & a_{79} \\ a_{81} & a_{82} & a_{83} & a_{84} & a_{85} & a_{86} & a_{87} & a_{88} & a_{89} \\ a_{91} & a_{92} & a_{93} & a_{94} & a_{95} & a_{96} & a_{97} & a_{98} & a_{99} \end{pmatrix}, \quad A = \begin{pmatrix} 011000000 \\ 000100000 \\ 000011000 \\ 000000000 \\ 000000110 \\ 000000001 \\ 000000000 \\ 000000000 \\ 000000000 \\ 000000000 \end{pmatrix} \quad (1)$$

Далее ГМ была передана в другую систему автоматизации, в результате имеются некоторые потери. Для наглядности описываем часть сборки переданной ГМ, представленной на рис. 3.

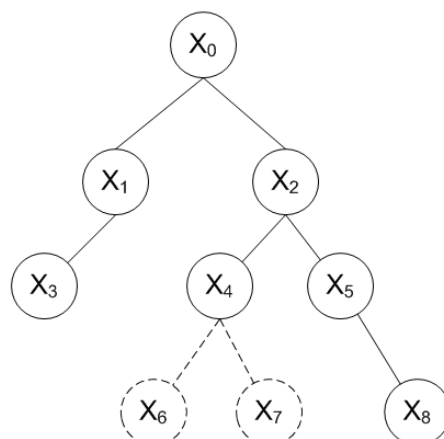


Рис. 3. Возможный результат конвертации ГМ

Матрица смежности переданной ГМ имеет тот же размер, что и исходная ГМ, так как переданная ГМ сравнивается с изначальной ГМ. Размер матрицы смежности определяется числом вершин графа, поэтому необходима процедура добавления нулевых строк и столбцов в «нужные места» (утраченные параметры), которые ещё потребуется определить. Матрица смежности и бинарные значения матрицы данной ГМ следующие

$$B = \begin{pmatrix} b_{11} & b_{12} & b_{13} & b_{14} & b_{15} & b_{16} & b_{17} & b_{18} & b_{19} \\ b_{21} & b_{22} & b_{23} & b_{24} & b_{25} & b_{26} & b_{27} & b_{28} & b_{29} \\ b_{31} & b_{32} & b_{33} & b_{34} & b_{35} & b_{36} & b_{37} & b_{38} & b_{39} \\ b_{41} & b_{42} & b_{43} & b_{44} & b_{45} & b_{46} & b_{47} & b_{48} & b_{49} \\ b_{51} & b_{52} & b_{53} & b_{54} & b_{55} & b_{56} & b_{57} & b_{58} & b_{59} \\ b_{61} & b_{62} & b_{63} & b_{64} & b_{65} & b_{66} & b_{67} & b_{68} & b_{69} \\ b_{71} & b_{72} & b_{73} & b_{74} & b_{75} & b_{76} & b_{77} & b_{78} & b_{79} \\ b_{81} & b_{82} & b_{83} & b_{84} & b_{85} & b_{86} & b_{87} & b_{88} & b_{89} \\ b_{91} & b_{92} & b_{93} & b_{94} & b_{95} & b_{96} & b_{97} & b_{98} & b_{99} \end{pmatrix}, \quad B = \begin{pmatrix} 011000000 \\ 000100000 \\ 000011000 \\ 000000000 \\ 000000000 \\ 000000000 \\ 000000001 \\ 000000000 \\ 000000000 \\ 000000000 \end{pmatrix} \quad (2)$$

Определение близости графов с применением метрик

Метрика близости графов рассчитывается на основании выражения расстояния Хэмминга, если требуется получить номинальные значения конвертации:

$$d_{ij} = \sum_{k=1}^p |x_{ik} - x_{jk}| \quad (3)$$

Получается следующее выражение:

$$d_{ij} = \sum_{i=1}^n \sum_{j=1}^n |a_{ij} - b_{ij}| \quad (4)$$

где a – параметры 1-й ГМ i -й строки и j -го столбца; b – параметры 2-й ГМ i -й строки и j -го столбца, полученной после конвертации; n – количество элементов.

Также метрика близости графов рассчитывается на основании меры Сёренсена, если требуется получить количественные значения конвертации:

$$K_S = \frac{2c}{a+b}, \quad (5)$$

где a – количество параметров 1-й ГМ, $a = \{X_1, X_2, X_3, X_4, X_5, X_6, X_7, X_8, X_9\}$, b – количество параметров на 2-й ГМ в результате конвертации, $b = \{X_1, X_2, X_3, X_4, X_5, X_8, X_9\}$, c – количество параметров, общих для 1-й и 2-й ГМ $c = \{X_1, X_2, X_3, X_4, X_5, X_8, X_9\}$.

Проведено исследование близости ГМ с применением теории графов; структура ГМ изделий и результаты передачи математически представлены в виде графа, содержат группы параметров: геометрия, атрибутивная информация, параметры файла, а также

параметры интеграции. Сравнения проводились на разных уровнях, в том числе в рамках сборки, в рамках отдельного уровня структуры, в рамках параметров узла.

Искажения для разных моделей очень часто не линейно зависят от количества элементов, а скорее от систем автоматизации проектирования и форматов. Поэтому требовалось выяснить, какие возможны потери при передаче в условиях мультивендорности, а также проведено тестирование для нейтральных форматов [7] программных сред. В результате экспериментов по конвертации с применением нейтральных форматов получались значения метрик на основании сравнения близости графов ГМ от 0 до 0,5. Для каждого случая и каждого формата это значение будет меняться, и необходимо каждый случай рассматривать детально: насколько полученный результат устроит, оценить какие были потери при конвертации, какие потребуются дополнительные затраты на восстановление. Если использовать системы автоматизации проектирования в рамках одного вендора, то выявлено, что проблемы при конвертации не зависят от форматов, а носят случайный характер. При применении разных вендоров в первую очередь выявлена зависимость от форматов и выявлен более высокий объем потерь при конвертации. Определено, что в определенных условиях возможно использовать нейтральные форматы, при условиях, что потери при конвертации параметров ГМ не повлияют на дальнейшую разработку и работу с изделием и не отразится негативным образом на дальнейшую поддержку этапов жизненного цикла изделий.

Заключение

Выявлено отсутствие поддержки конвертации данных между системами автоматизации проектирования, в форме обобщенной машинно-независимой модели на основании сравнения близости графов. Модель выявляет позволяет оценить затраты, необходимые на восстановление данных. Предложена расширенная модель структуры параметров с применением графов. Определены проблемы интеграции систем автоматизации проектирования. Предложена оценка конвертации ГМ между системами автоматизации в рамках ЕИП на основании метрик и определение значений.

Библиографический список

1. Методы проектирования информационно-управляющих и телекоммуникационных систем / В.Г. Баранов, Л.М. Вдовин, Т.И. Горячева [и др.]; под ред. В.Р. Милова, В.Г. Баранова. М.: Радиотехника, 2016. 216 с.
2. Яблочников Е.И., Фомина Ю.Н., Саломатина А.А. Компьютерные технологии в жизненном цикле изделия: учебное пособие. СПб.: СПбГУ ИТМО, 2010. 180 с.
3. O. Shilolitsky. How to re-invent CAD / PDM integration? / [Электронный ресурс]. — Режим доступа: URL: <http://beyondplm.com/2014/05/19/how-to-re-invent-cad-pdm-integration/> (дата обращения: 15.01.2021).
4. Татт У. Теория графов: Пер. с англ. М.: Мир, 1988. 424 с.
5. Филинских А.Д., Соснина О.А., Бойтяков А.А. Иерархическое пространство параметров геометрической модели // Вестник БГТУ им. В.Г. Шухова. 2015. №2. С.131-134.
6. Мерзляков И.Н., Филинских А.Д., Райкин Л.И., Бойтяков А.А. Модель передачи данных на основании структуры параметров геометрических моделей с применением фреймов // Информационно-измерительные и управляющие системы. 2017. №8. С.34-41.
7. Филинских А.Д., Мерзляков И.Н. Управление проектными данными: [учебное пособие для направления 09.04.02 «Информационные системы и технологии»]. Н. Новгород: НГТУ, 2017. 92 с.

A.A. Boytyakov

PRINCIPLES SUPPORTING THE PROCESS OF DATA CONVERTING GEOMETRIC MODEL IN DISSIMILAR AUTOMATED SYSTEMS USING A COMPARISON OF THE PROXIMITY OF THE STRUCTURES OF GEOMETRIC MODELS

Nizhny Novgorod state technical university n.a. R.E. Alekseev

The paper contains a method for comparing the proximity of structures of geometric models in dissimilar automated systems using graphs, as part of supporting the conversion process between systems using the example of CAD and PDM systems. Industrial enterprises competitiveness depends heavily on the qualitative interaction between CAD and PDM systems, the efficiency of work management at the design stage, and their successful integration into the common information space structure.

Keywords: Geometrical model, data transfer format, converting parameters, CAD system, PDM system.

ИССЛЕДОВАНИЕ ФУНКЦИОНАЛЬНЫХ ВОЗМОЖНОСТЕЙ УДАЛЕННОГО УПРАВЛЕНИЯ КАМЕРОЙ МОБИЛЬНОГО УСТРОЙСТВА ДЛЯ ФОТОГРАММЕТРИЧЕСКОЙ СЪЕМКИ

Нижегородский государственный технический университет им. Р.Е. Алексеева

Рассматриваются основные принципы работы и возможности приложений для IP-видеонаблюдения для задач фотограмметрии. Исследование проводится в рамках работы по автоматизации процесса получения входных данных. Проанализированы возможности удаленного доступа к камере снимающего устройства, а также метод загрузки полученных фотографий на персональный компьютер с установленным ПО для их обработки.

Ключевые слова: фотограмметрия; 3D-модель; удаленный доступ; IP-видеонаблюдение.

Визуальное изображение настолько прочно вошло в нашу жизнь, что уже сложно представить без 3D-моделирования строительство [1], индустрию кино, игр и мобильных приложений [2]. Примечательно, что с развитием и популяризацией 3D-моделирования увеличивается число методов создания визуальной модели объекта [3]. Одной из относительно простых в реализации технологий, обеспечивающих высокое качество 3D-модели, является технология фотограмметрии - получения цифровой модели объекта по серии его снимков

Важной задачей фотограмметрии является получение корректных входных данных (фотографий) для систем их последующей обработки. Сегодня это обеспечивается путем внедрения в процесс сканирования дорогостоящего профессионального фотограмметрического оборудования и/или наличия большого количества цифровых камер [4].

Многие программы могут обрабатывать фотографии, снятые любой цифровой камерой: от камеры современного смартфона до профессионального фотограмметрического оборудования. Помимо того, что передача информации из одной среды в другую изначально может приводить к потерям данных [5], большую роль играет и условия получения входных данных. Человеческий фактор получения исходных изображений зачастую приводит к их искажению. Это связано с несоблюдением ряда требований к проведению съемки объекта/сцены (процентное соотношение перекрытия сцены, взаимное расположение объектов в процессе съемки, попадание в кадр нежелательных объектов на передний план и многое другое).

Для устранения этих факторов, влияющих на точность входных изображений, предлагается минимизировать участие человека в процессе съемки путем автоматизации данного этапа. С этой целью предлагается исследовать возможность удаленного управления снимающим устройством, а также выгрузки полученных изображений на персональный компьютер для их дальнейшей обработки.

Автоматизация процесса получения исходных фотографий объекта для последующей обработки в фотограмметрических системах и получения итоговой 3D-модели объекта предполагает решение проблемы выбора способа удаленного управления снимающим устройством.

Принципиальная схема получения 2D-снимков объекта при минимальном участии человека представлена на рис. 1.

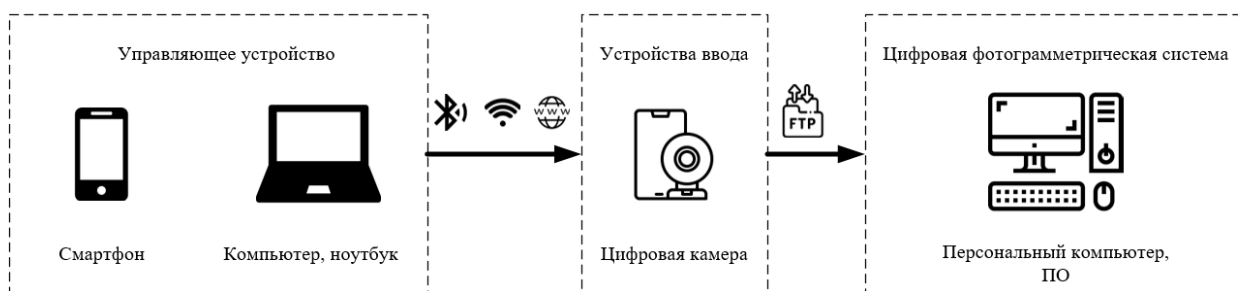


Рис. 1. Обобщенная принципиальная схема получения исходных изображений

В качестве снимающего устройства предлагается использовать смартфон, управление камерой которого осуществляется удаленно. В качестве управляющего устройства при этом можно использовать другое мобильное устройство или компьютер. Подключение между ними может происходить посредством Bluetooth или Wi-Fi (устройства должны находиться в одной локальной сети), а также через сеть Интернет. Для реализации предлагается использовать технологию p2p (англ. peer-to-peer – «узел к узлу») — разделение вычислительных ресурсов и служб посредством прямого обмена ресурсами между участниками сети [6]. Сети p2p ориентированы на специфические приложения обмена файлами и взаимодействия пользователей, в том числе активно применяются в системах IP-видеонаблюдения.

Рассмотрим самый простой вариант – управление камерой смартфона через приложения IP-видеонаблюдения [7].

В работе использовался смартфон Huawei Honor 10 (Android 10, процессор HiSilicon Kirin 970, 4x1.8 + 4x2.36 ГГц, видеоядро Mali-G72 MP12, 4 Гб ОЗУ, Wi-Fi IEEE 802.11ac, IEEE 802.11b, IEEE 802.11a, IEEE 802.11g, IEEE 802.11n, 24+16 Мп) и приложение IPWebcam [8].

Выбор мобильного приложения обоснован функциональными возможностями, простотой в настройке и дружелюбным интерфейсом, что подтверждается высокими оценками и положительными отзывами в PlayMarket (табл. 1).

Таблица 1

Характеристики IPWebcam

Оценка в Google Play	4,0
Количество загрузок в Play Маркет	10 000 000+
Страна-разработчик	Россия
Совместимость приложения с другими версиями ОС Android	Зависит от устройства
Размер дистрибутива	21Мб
Версия приложения	1.15.0г.769 (aarch64) от 29.10.20
Покупки в приложении	От 299 до 4990 руб. за единицу
Интерфейс	Настраиваемый интерфейс камеры
Возможность самостоятельной настройки	Присутствует
Максимальная нагрузка на систему (CPU/RAM) в состоянии покоя	0.0% (76 Мбайт)
Максимальная нагрузка на систему (CPU/RAM), локальная сеть	0.0% (67 Мбайт)
Максимальная нагрузка на систему (CPU/RAM), интернет-вещание	1,9% (76-200 Мбайт)
Расход аккумулятора	5,3%
Стабильность	Без сбоев
Трансляция видео по сети	Высокое качество
Трансляция видео удаленно через Интернет	Высокое качество

Разрешение фото/ видео	1920x1080
Ручная настройка битрейта, формата файла и трансляции	Присутствует
Архив	Присутствует
Маскировка (шпионская камера)	Присутствует
Ночной режим	Присутствует
Работа с облачными сервисами	Присутствует

В качестве объекта съемки был выбран глобус – как простая фигура оптимального размера.

Устройства были подключены к домашней локальной сети. Управление происходило через web-интерфейс (рис. 2).

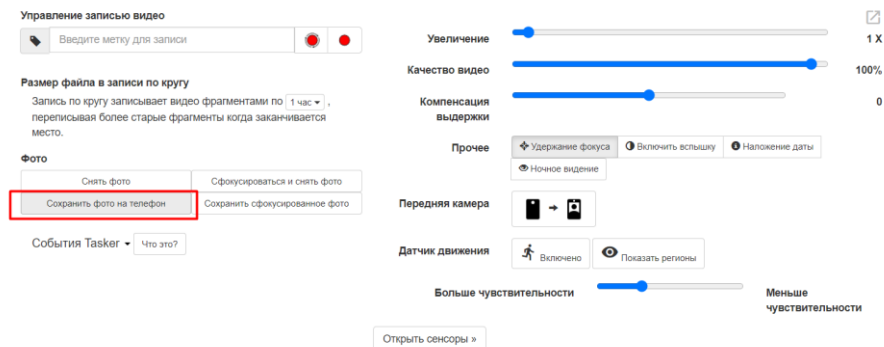


Рис. 2. Web-интерфейс IPWebcam

Web-интерфейс IPWebcam позволяет управлять фокусом, зумированием, а также световыми настройками сцены.

В ходе съемки также использовалась кольцевая LED лампа 26 см, штатив 2,1 метра, держатель для телефона RING FILL LIGHT. Условия освещенности специально были заданы как неидеальные (искусственно затемненные), чтобы проанализировать возможность использования рассеянного света от кольцевой лампы (рис. 3).

От грамотно выставленного освещения зависит не только полнота данных, считываемых с исходных изображений. 3D-модель объекта, полученная с помощью технологии фотограмметрии, уже обладает хорошими «световыми» данными, что избавляет от необходимости настраивать глобальное освещение итоговой сцены при 3D-моделировании [9].



Рис. 3. Сцена съемки

Характеристики сцены:

- расстояние до объекта 52 см (от камеры до центра объекта);
- высота источника света (у основания) 88 см;
- высота камеры 92 см.

В среднем отклик системы для создания снимка составил 1,30 с, при изменении сцены – 2 с (смена ракурса, регулирование положения телефона относительно сцены).

Всего было сделано 48 фотографий в разрешении 1920x1080 пикселей. Это максимальное разрешение для бесплатной версии приложения.

Передача изображений для дальнейшей обработки фотограмметрических программах была организована через FTP-сервер.

Снимающее устройство Honor 10, на котором были сохранены снимки, выступало как сервер (рис. 4), а клиент был организован на ноутбуке с установленным фотограмметрическим ПО через приложение FileZilla [10] - рис. 5.

Выгрузка фотографий на ноутбук совершалась после завершения съемки. Данный способ не предусматривал задачи загрузки фотографий по мере их добавления в архив приложения и/или хранилище телефона.

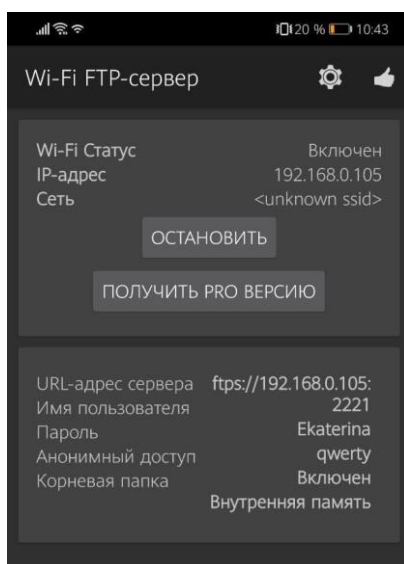


Рис. 4. Интерфейс FTP-сервера на мобильном устройстве

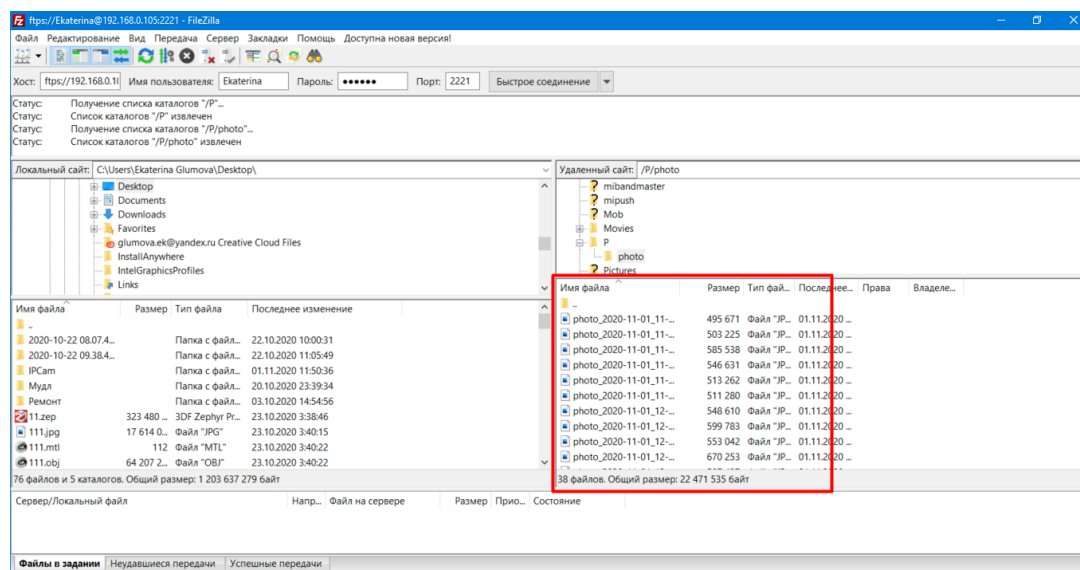


Рис. 5. Интерфейс FTP-клиента на компьютере

Загрузка 48 фотографий заняла 48 с – по 1 с на каждое изображение (рис. 6).

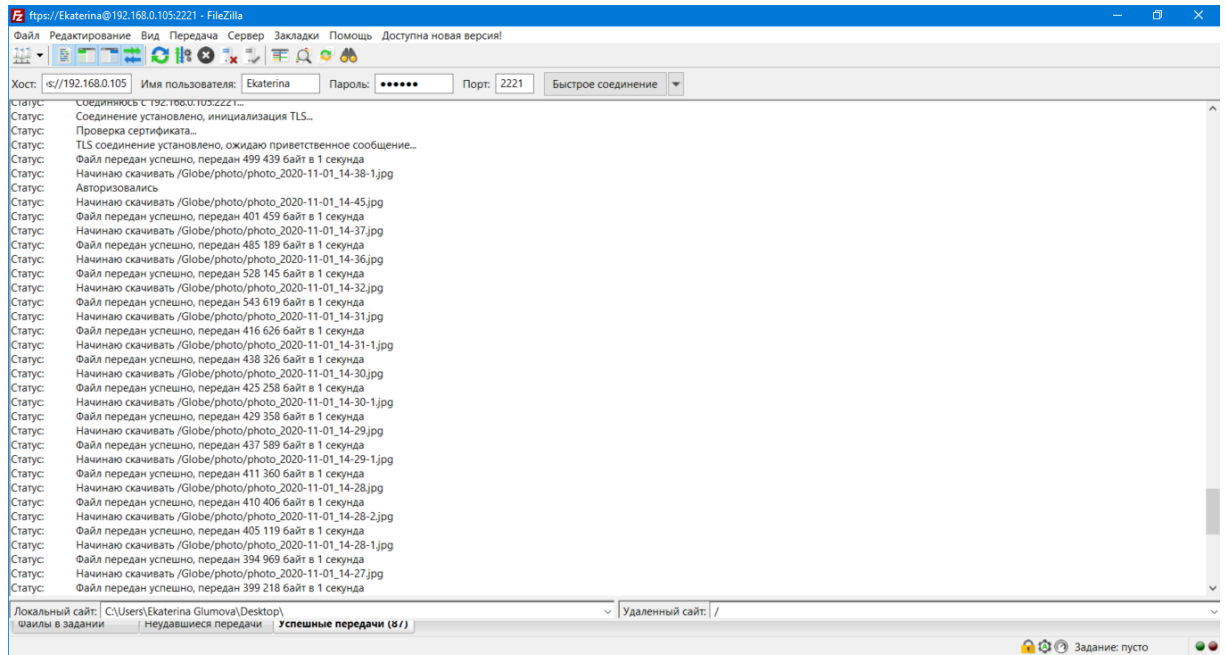


Рис. 6. Статусная строка FTP-клиента при загрузке файлов

Далее все фотографии были перенесены в проекты фотограмметрического ПО (Regard3D [11] и 3DFZephyrFree [12]) для дальнейшей обработки.

Реконструкция объекта глобуса производилась на ноутбуке Lenovo Ideapad 330 15IKB, Intel Core i5 7200U 2,50 ГГц, x64, 12,0 Гб ОЗУ, Видеопроцессор AMD Radeon 530, GDDR5, 2 Гб.

Результат 3d-реконструкции в программе 3DFZephyrFree представлен на рис. 7. Модель получилась без серьезных разрывов, но с большими искажениями, и требует доработки в сторонних программах. Необходимо отметить, что лишнюю геометрию в виде окружения можно убрать непосредственно перед экспортом в этой системе.

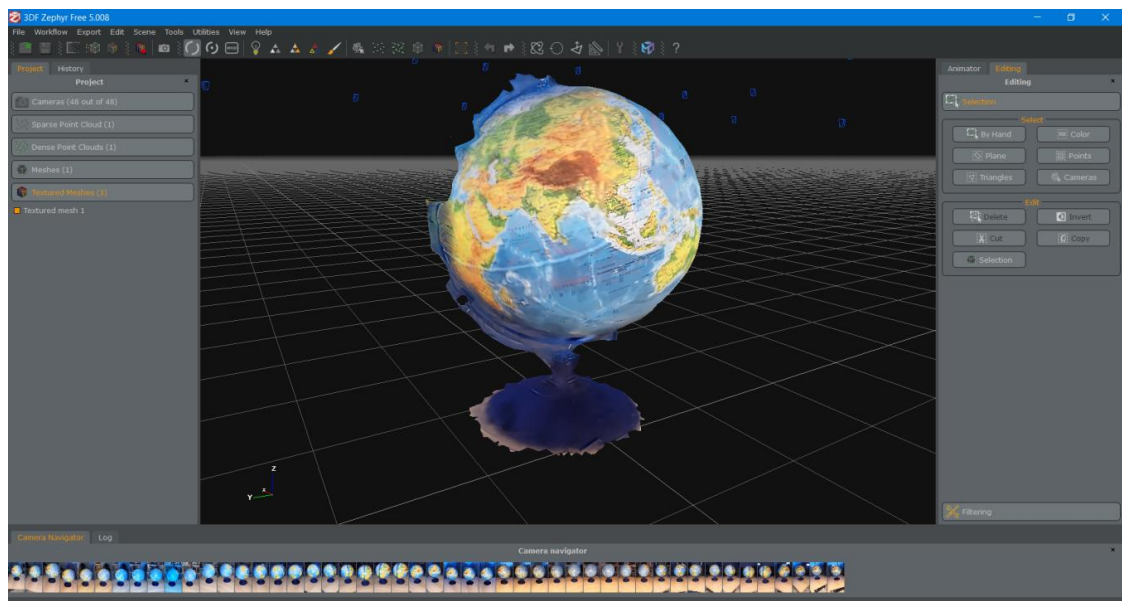


Рис. 7. Результат 3D-реконструкции в программе 3DFZephyrFree

Всего было создано 10 495 вершин, общее время реконструкции составило 4,88 мин.

Результат 3D-реконструкции в программе Regard3D представлен на рис. 8. Модель получилась более детализированная (видны все надписи на объекте), нет искажений из-за бликов от источника света. Но на модели отсутствуют такие серьезные части, как подставка и крепежный элемент.

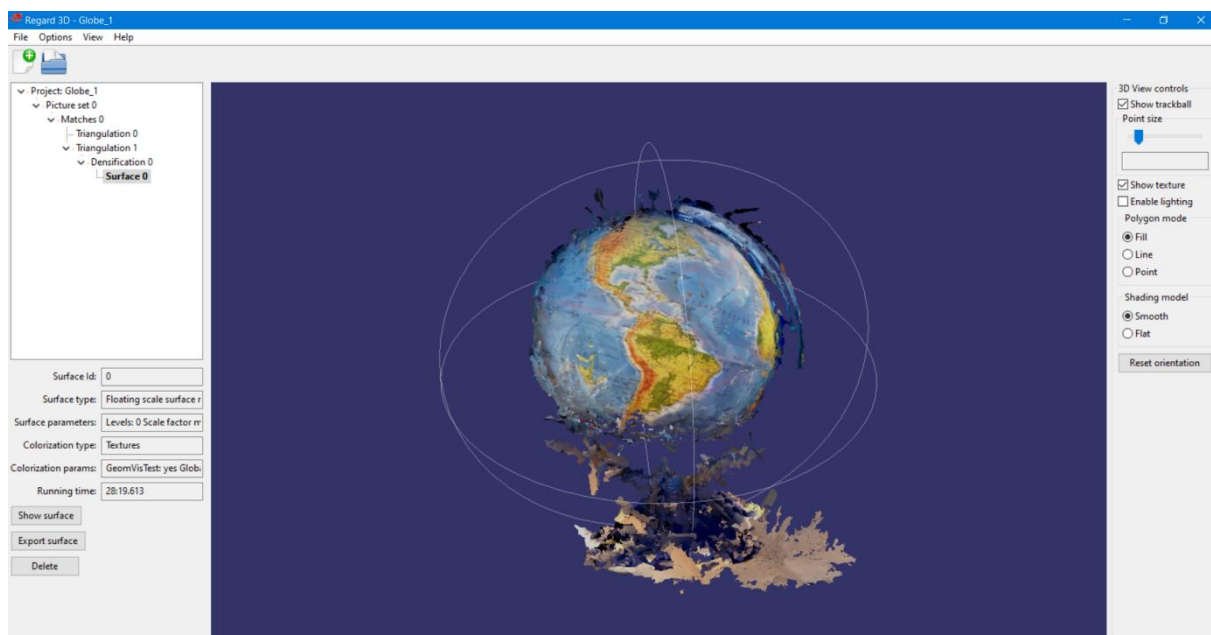


Рис. 8. Результат 3d-реконструкции в программе Regard3D

Всего было создано 320633 вершины, общее время реконструкции составило 126,74 мин ($\approx 2,2$ ч).

Как видим, ни одну из моделей нельзя использовать как готовый 3D-объект. Следовательно, работу с IP-камерой можно применять, когда не требуется большого процента детализации, а нужна лишь форма объекта. Также этот метод может найти применение для получения материалов, которые впоследствии будут использоваться в создании текстурных карт.

Кольцевую лампу можно использовать в качестве источника рассеянного света, так как на исходных фотографиях не были отмечены тени и блики, а цвета и форма объекта переданы достоверно.

Способ передачи изображений через FTP-сервер никак не замедляет и не искажает работу. Изображения были переданы без сжатия, данные о камере на них сохранились. Значит, такой вариант работы возможен.

Приложения IP-видеонаблюдения могут обеспечить приемлемое качество фотографий и видео, отвечающее задачам фотограмметрии. Поскольку использовалась лишь бесплатная версия приложения, ограничивающая разрешение исходных фотоизображений значением 1920x1080 пикселей, не удалось получить более качественный результат. Но даже использование бесплатных программных продуктов и непрофессионального оборудования съемки позволяет получать 3D-модель объекта приемлемого качества.

Библиографический список

1. Соснина О.А., Филинских А.Д., Ложкина Н.А. Анализ методов создания виртуальных моделей нетривиальных форм // Информационные технологии. – 2019. – Т. 25. – № 11. – С. 679-681.
2. Соснина О.А., Филинских А.Д., Коротаева А.С. Сравнение способов создания низкополигональной 3D-модели // Информационные технологии. – 2017. – Т. 23. – № 8. – С. 564-568.

3. Малышева А.И., Томчинская Т.Н. Особенности низкополигонального моделирования и текстурирования в мобильных приложениях //КОГРАФ-2019. – 2019. – С. 51-54.
4. Glumova E.S., Filinskikh A.D., Investigation of algorithms for generating surfaces of 3D models based on an unstructured point cloud. (2020) CEUR Workshop Proceedings, 2763, pp. 243-249. URL: <https://www.scopus.com/record/display.uri?eid=2-s2.0-85097558020&origin=inward&txGid=df4c0a1dcd78a5cab532af67b8b36247>
5. Филинских А.Д. Информационная метрика передачи и восстановления геометрических моделей в профессиональных программных средах: дис. – Нижегородский государственный технический университет им. РЕ Алексеева, 2013.
6. Радченко Г.И. Распределенные вычислительные системы //Челябинск: Фотохудожник. – 2012. – С. 184.
7. IP-камера: что это такое, как работает, какие бывают IP-камеры [Электронный ресурс] URL: <https://www.dssl.ru/publications/stati/ip-kamera-chto-eto-takoe-kak-rabotaet-kakie-byvayut-ip-kamery/>
8. IPWebcam [Электронный ресурс] URL: <https://play.google.com/store/apps/details?id=com.pas.webcam&hl=ru&gl=US>
9. Серова М.А., Филинских А.Д. Актуальные методы имитации глобального освещения в 3d-моделировании //КОГРАФ-2020. – 2020. – С. 143-
10. 148FileZilla FTP сервер [Электронный ресурс] URL:https://filezilla.ru/documentation/FileZilla_FTP_Server
11. Regard3D[Электронный ресурс] URL:<https://www.regard3d.org/>
12. 3DFZephyrFree [Электронный ресурс]URL: <https://www.3dflow.net/3df-zephyr-free/>

E.S. Glumova, A.D. Filinskikh

INVESTIGATION OF MOBILE DEVICE REMOTE CAMERA CONTROL FUNCTIONALITY FOR PHOTOGRAMMETRIC SURVEYING

Nizhniy Novgorod state technical university n.a. R.E. Alekseev

Basic principles and capabilities of IP video surveillance applications for photogrammetry tasks are considered. The research is carried out within the framework of the work on automation of the process of receiving input data. Opportunities of remote access to the camera of a shooting device are analysed as well as methods of downloading the received photos to a personal computer with installed software for their processing. Recommendations on the use of the software are given.

Keywords: photogrammetry; 3D-model; remote access; IP-video monitoring.

ФОРМИРОВАНИЕ ГРУППЫ ЭКСПЕРТОВ ПРИ ЭКСПЕРТНО-ЭВРИСТИЧЕСКОМ МЕТОДЕ ОЦЕНКИ ИНТЕРФЕЙСОВ

Нижегородский государственный технический университет им. Р.Е. Алексеева

Исследование направлено на формирование нового метода оценки интерфейсов, обеспечивающего многокритериальность и устраняющую недостатки предшествующих методов. Предлагается совмещение экспертного и эвристического подхода, что позволяет выявлять широкий спектр проблем UI/UX, обеспечивает компетентность оценки и снижает уровень недоверия к эксперту. Затем описывается исследовательский пример формирования весового коэффициента эксперта. Практическим путем удалось доказать преимущество использования нового подхода к оценке эксперта для нахождения его весового коэффициента за счет включения в результирующую оценку трех составляющих, подтверждающих его компетентность; проработанной системы вопросов для анализа профессиональных навыков, учитывающей наиболее значимые аспекты для проведения оценки интерфейсов; ранжированной системы учета источников аргументации. Настоящее исследование служит отправной точкой для дальнейших исследований области оценки интерфейсов. Кроме того, метод позволяет устранить такие недостатки использования системы ГОСТов как отсутствие количественных составляющих и снижают уровень недоверия к эксперту. Заметим, что разработанный метод позволяет с достаточно высокой точностью провести оценку пользовательских интерфейсов также благодаря проработанной системе эвристик, учитывающих как общие, так и специфические особенности тех или иных групп интерфейсов.

Ключевые слова: экспертная оценка, эвристическая оценка, методы оценки, пользовательский интерфейс, весовой коэффициент эксперта.

Введение

Стремительное развитие технологий определяет появление новых типов интерфейсов с разнообразными возможностями, характеристиками и типами управления, например, бесконтактное управление на основе жестов или голоса. Прогресс в области UI в совокупности с изменчивостью требований создают сложности для стандартизации и объективной оценки качества юзабилити инновационных, т.е. обеспечивающих интуитивный подход – научение пользователя на основе предыдущего опыта взаимодействия с различными интерфейсами, типов человеко-компьютерного взаимодействия.

В современных условиях дизайн интерфейсов имеет ключевое значение в процессе взаимодействия пользователя и устройства. Детальная проработка всех аспектов пользовательского интерфейса улучшает его качество и повышает скорость взаимодействия человека и машины. Дополнительно сокращается себестоимость продукта, временные и ресурсные затраты за счет устранения большинства ошибок еще на этапе прототипирования и, как следствие, сокращения числа итераций переработки проекта при использовании любой модели ЖЦ. Также облегчается понимание и научение при работе с UI [1]. В связи с этим возрастает значимость своевременной многокритериальной оценки качества интерфейса для обеспечения потенциального успеха разрабатываемого программного продукта.

Предшествующие альтернативные методы оценки интерфейсов

Существует несколько способов оценки качества пользовательского интерфейса с помощью межгосударственных стандартов: ГОСТ Р ИСО 14915–1–2010 «Эргономика мультимедийных пользовательских интерфейсов» [2], ГОСТ Р ИСО 9241–210–2012 «Эргономика взаимодействия человек-система» [3], ГОСТ 28195–89 «Оценка качества

программных средств» [4], которые представляют собой список требований и рекомендаций по проектированию и функциональным особенностям различных элементов UI с примерами. Однако данные стандарты не учитывают весь спектр параметров эффективности и результативности интерфейсов. Их существенным недостатком является отсутствие количественной составляющей, что ведет к невозможности их формализации. Стремительное развитие технологий и смена тенденций дизайна ведет к быстрому устареванию данных способов оценки и появлению необходимости формирования новых критериев оценки качества интерфейсов. Таким образом, становится невозможным применение государственных стандартов качества юзабилити для оценки современных прототипов интерфейсов.

Альтернативными вариантами оценки UI являются использование эвристик и проведение экспертизы. Метод эвристической оценки привлекателен своей простотой и прозрачностью: небольшая группа людей анализирует интерфейс, основываясь на общепризнанных принципах юзабилити, впервые разработанных Якобом Нильсеном совместно с Рольфом Моличем [5].

Достоинством данного метода является потенциальная возможность формализации, недостатком – невозможность формирования полного списка эвристик ввиду динамичности развития интерфейсов и технологий в целом. Однако следует отметить, что четкий контрольный список (наиболее значимые эвристики) может использоваться даже непрофессионалами в проектировании интерфейсов, что позволяет проводить оценку обывателями - представителями целевой аудитории, что, в свою очередь, позволяет узнать об их предпочтениях.

В вопросе альтернативных методов конкуренцию эвристической оценке может составить экспертиза. Экспертная проверка на соответствие стандартам проектирования интерфейса проводится специалистами в области взаимодействия человека и компьютера или специалистами по человеческому фактору и позволяет быстро и качественно распознать проблемы юзабилити. Однако зачастую, в отличие от простых обывателей, эксперты мыслят шаблонно и не способны увидеть нестандартные проблемы и их оригинальные решения. Существенным недостатком метода также являются вопросы об уровне доверия к эксперту и степени субъективности оценок.

Проблематика исследования

Исходя из вышеизложенного, можно сделать вывод о несостоятельности существующих методов в условиях динамично развивающихся интерфейсов. Возникает необходимость создания нового метода, обладающего следующими свойствами:

- гибкость, исходя из функциональных особенностей, сложности и/или сферы применения интерфейса;
- быстрота и простота применения метода;
- потенциальная возможность формализации;
- возможность уменьшения или полного элиминирования субъективного восприятия.

Предлагаемый новый метод проведения оценки интерфейсов

В качестве нового метода проведения оценки предлагается алгоритм, соединяющий достоинства экспертизы и эвристического подхода. Он заключается в следующем:

1. Каждый эксперт имеет собственный весовой коэффициент, формирующийся с учетом его профессиональных компетенций, уровня погруженности в оцениваемую сферу, а также факторов, влияющих на степень аргументированности.
2. Эвристики, в свою очередь, сформированы и разделены на категории таким образом, чтобы максимально охватывать как общие принципы юзабилити, так и особенности интерфейсов конкретной сферы бизнеса.

3. Помимо всего прочего, общая оценка интерфейса формируется с учетом важности критерия в общей совокупности эвристик.

Для апробации разработанного метода было проведено два эксперимента по оценке мобильных интерфейсов, распределенных по группам в зависимости от сферы бизнеса или области деятельности (социальные, коммерция, организация досуга и труда и др.). Каждый опыт включал в себя несколько этапов (рис. 1):



Рис. 1. Этапы проведения экспериментов

Апробация разработанного метода. Определение весового коэффициента эксперта

Первый этап эксперимента включал в себя формирование весового коэффициента эксперта. Особенностью данной метода является то, что анкетирование, а затем и обработка результатов, проводятся с учетом показателей компетентности и весомости привлекаемых экспертов.

Обычно возникают большие затруднения при определении степени объективности весовых коэффициентов. Помимо этого, отмечаются недостатки, связанные с невысокой степенью обоснованности и, соответственно, низкой степенью доверия к получаемым результатам. С целью повышения объективности и степени доверия к получаемым в ходе экспертного опроса значениям весовых коэффициентов предлагается методика определения их величин, которая учитывает уровень компетентности привлекаемых экспертов, определенный на основе их самооценки и оценки степени их знакомства с областью знаний. С этой целью были разработаны специальные анкеты (рис. 2):

Рис. 2. Форма для анкетирования экспертов

Подобная методика впервые была рассмотрена Г. М. Добровым и его коллегами [6], занимающимися вопросами анализа компетентности эксперта.

В первом пункте анкеты эксперту предлагалось ответить на вопросы для определения базового весового коэффициента (каждый пункт дает от 0,1 балла).

1. Профильное образование в сфере проведения оценок (0,1 балла)
2. Опыт проведения оценочной экспертизы (0,1 балла)
3. Наличие публикаций по теме экспертных оценок (0,1 балла)
4. Профильное образование в сфере разработки интерфейсов (0,1 балла)
5. Опыт разработки интерфейсов (0,2 балла)
6. Наличие публикаций по теме разработки интерфейсов (0,2 балла)
7. Работа в профильной организации по разработке интерфейсов (0,2 балла)

Критерии выбраны таким образом, чтобы выявить степень знакомства предполагаемого эксперта с областью исследования. Наиболее значимыми критериями являются «Опыт разработки интерфейсов» и «Работа в профильной организации по разработке интерфейсов», т.к. только наличие непосредственного практического опыта обеспечивает полное понимание не только основ, но и специфики разработки различных интерфейсов.

Во втором пункте анкеты эксперту предлагалось оценить показатель степени личной уверенности в соответствии собственных суждений действительности. Согласно исследованиям ведущих специалистов в области экспертных оценок, описанных во многих учебных пособиях [7], установлено, что точность усредненной экспертной оценки тем выше, чем выше среднее значение самооценки членов группы (рис. 3). Поэтому в состав экспертной группы рекомендуется брать специалистов с высокими значениями самооценки. В расчетах показатель степени уверенности использовался для определения весомости каждого эксперта.

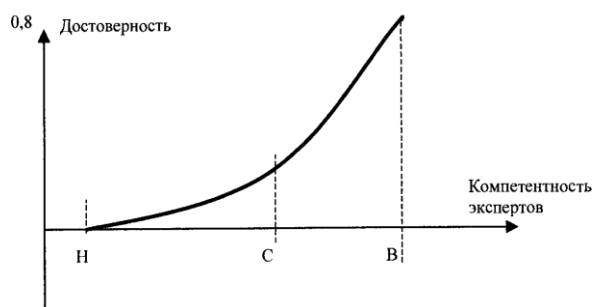


Рис. 3. Зависимость достоверности экспертного оценивания от уровня компетентности экспертов (по их самооценке)

В третьем пункте анкеты осуществлялась самооценка степени влияния различных источников аргументации на ответ эксперта. В частности, предусматриваются следующие критерии: проведенные теоретические и практические исследования по направлению; непосредственное участие в проведении подобных типов экспертизы; практический опыт как специалиста по данному вопросу и (или) участие в разработке, интуиция эксперта.

Весомость j -го эксперта определяется в соответствии с выражением [8]:

$$w_j = k_j p_j \quad (1)$$

где p_j – степень уверенности в прогнозе j -го эксперта, которая определяется нормированным значением соответствующей оценки, поставленной экспертом в п.2 (умножением ее на 0,01);

k_j – компетентность j -го эксперта, $j=1, \dots, n$.

Компетентность j -го эксперта в выражении (1) может быть определена структурой аргументов, послуживших ему основанием для ответа, а также степенью знакомства с рассматриваемым вопросом и рассчитывается по выражению:

$$k_j = \frac{k_{ja} + k_{jb}}{2}, \quad (2)$$

где k_{j_a} – коэффициент аргументированности j -го эксперта, который определяется путем наложения численных значений таблицы 1 на клетки аналогичной таблицы, заполненной экспертом, и суммирования соответствующих численных значений таблицы 1;
 k_{j_s} – коэффициент степени знакомства j -го эксперта, который определяется суммарным значением оценки, формируемой в результате прохождения анкеты (п.1).

Таблица 1

Степень влияния источника аргументации

Источник аргументации	Степень влияния источника документации на Ваше мнение		
	Высокая	Средняя	Низкая
Проведенные теоретические и практические исследования по направлению	0,4	0,3	0,2
непосредственное участие в проведении подобных типов экспертизы	0,3	0,2	0,1
практический опыт как специалиста по данному вопросу и(или) участие в разработке	0,2	0,1	0
Ваша интуиция	0,1	0	0

Значения таблицы распределены таким образом, чтобы суммарно анкетированный мог получить 1 балл.

В п. 1 анкетированный также может набрать не более 1 балла. Результирующий коэффициент компетентности получается путем вычисления среднего арифметического коэффициента аргументированности и коэффициента значимости. Таким образом, он не может превышать 1, т.е. 100%.

Данная методика позволяет выступать в качестве эксперта любому рядовому пользователю, даже никогда не работавшему с интерфейсами, однако степень влияния его оценочных суждений в формировании общей оценки будет низкой.

Результаты расчетов весовых коэффициентов для первого эксперимента представлены в таблице 2:

Таблица 2

Весовые коэффициенты экспертов

	w_j	% в общей оценке
Эксперт 1	0,25	6,693440428
Эксперт 2	0,28	7,49665328
Эксперт 3	0,385	10,30789826
Эксперт 4	0,15	4,016064257
Эксперт 5	0,49	13,11914324
Эксперт 6	0,315	8,43373494
Эксперт 7	0,24	6,425702811
Эксперт 8	0,315	8,43373494
Эксперт 9	0,28	7,49665328
Эксперт 10	0,2	5,354752343
Эксперт 11	0,35	9,3708166
Эксперт 12	0,3	8,032128514
Эксперт 13	0,18	4,819277108

Выводы, дальнейшие пути исследования и развития метода

Таким образом, практическим путем удалось доказать преимущество использования нового подхода к оценке эксперта для нахождения его весового коэффициента за счет:

- включения в результирующую оценку трех составляющих – профессиональных навыков и компетенций, самооценки эксперта на основе его вовлеченности в сферу интерфейсов и аргументированности мнения эксперта;
- проработанной системы вопросов для анализа профессиональных навыков, учитывающей наиболее значимые аспекты для проведения оценки интерфейсов;
- ранжированной системы учета источников аргументации.

Данные аспекты позволяют устранить такие недостатки использования системы ГОСТов как отсутствие количественных составляющих и снижают уровень недоверия к эксперту как при использовании экспертного подхода в его первоначальном виде.

Хотелось бы отметить, что разработанный метод позволяет с достаточно высокой точностью провести оценку пользовательских интерфейсов также благодаря проработанной системе эвристик, учитывающих как общие, так и специфические особенностей тех или иных групп интерфейсов. Это в полной мере подтверждается анализом результатов проведенных вычислительных экспериментов, которым будут посвящены дальнейшие работы.

Библиографический список

1. Зенг В., «Оценка качества проектирования пользовательских интерфейсов нового поколения», Известия ТулГУ. Технические науки, № 12, 2019, С. 404-410.
2. «ГОСТ Р ИСО 14915-1-2010. Эргономика мультимедийных пользовательских интерфейсов.».
3. «ГОСТ Р ИСО 9241-210-2016 Эргономика взаимодействия человек-система. Часть 210. Человеко-ориентированное проектирование интерактивных систем: дата введения 2017-12-01.».
4. «ГОСТ 28195-89 Оценка качества программных средств. Общие положения: дата введения».
5. «Heuristic evaluation,» Wikimedia Foundation, Inc., 2019.
6. Добров Г. М., Ершов Ю. В., Левин Е. И. и Смирнов Л. П., Экспертные оценки в научно-техническом прогнозировании, В. С. Михалевич, Ред., Киев: Наук. думка, 1974, р. 160.
7. Шишкова Г., Менеджмент (Управленческие решения): учебно-методический модуль, Москва: Издательство Ипполитова, 2002.
8. Глушков В., Методика программного прогнозирования развития науки и техники, Москва: Гос. Комитет Сов. Мин. СССР по науке и техн., 1971, с. 270.
9. Marcio Silva C., Macedo V., Lemos R., Okimoto M.L.L.R. (2014) Evaluating Quality and Usability of the User Interface: A Practical Study on Comparing Methods with and without Users. In: Marcus A. (eds) Design, User Experience, and Usability. Theories, Methods, and Tools for Designing the User Experience. DUXU 2014. Lecture Notes in Computer Science, vol 8517. Springer, Cham. https://doi.org/10.1007/978-3-319-07668-3_31
10. Robin Jeffries, James R. Miller, Cathleen Wharton, Kathy M. Uyeda, User Interface Evaluation in the Real World: A Comparison of Four Techniques. In: Software and Systems Laboratory, HPL-91-03, January, 1991
11. Шевелева М.М., Зарубин И.Б., Гуляева У.И., Сорокин С.Б., Лабаев А.М., «КОЛИЧЕСТВЕННАЯ ОЦЕНКА КОРРЕКТНОСТИ САЙТОВ ВЫСШИХ УЧЕБНЫХ ЗАВЕДЕНИЙ», В сборнике: КОГРАФ - 2020. сборник материалов 30-й Всероссийской научно-практической конференции по графическим информационным технологиям и системам. Нижний Новгород, 2020. С. 368-372.

12. Шевелева М.М., Зарубин И.Б., Гуляева У.И., Сорокин С.Б., Лабаев А.М., “КРИТЕРИИ ОЦЕНКИ УДОБСТВА ИСПОЛЬЗОВАНИЯ САЙТОВ ВЫСШИХ УЧЕБНЫХ ЗАВЕДЕНИЙ”, В сборнике: Информационные системы и технологии ИСТ-2020. Сборник материалов XXVI Международной научно-технической конференции. Нижегородский государственный технический университет им. Р.Е. Алексеева. 2020. С. 392-398.

U.I. Gulyaeva

FORMATION OF A GROUP OF EXPERTS USING THE EXPERT-HEURISTIC METHOD FOR EVALUATING INTERFACES

Nizhny Novgorod state technical university named after R. E. Alekseev

Purpose: The research is aimed at forming a new method for evaluating interfaces that provides multi-criteria and eliminates the shortcomings of previous methods.

Design/methodology/approach: A combination of an expert and a heuristic approach is proposed, which allows identifying a wide range of UI/UX problems, ensures the competence of the assessment and reduces the level of distrust in the expert. Then the research example of forming the expert's weighting factor is described.

Findings: A practical way managed to prove the advantage of using a new approach to assessment to find its weight by the inclusion in the resulting estimation of the three components, confirming its competence; developed a system of questions for review of skills, taking into account the most significant aspects for the evaluation of interfaces; a ranked system of accounting for sources of argument.

Research limitations/implications: This study serves as a starting point for further research into the field of interface evaluation.

Originality/value: In addition, the method eliminates such disadvantages of using the GOST system as the lack of quantitative components and reduces the level of distrust of the expert.

Note that the developed method allows you to evaluate user interfaces with a fairly high accuracy also thanks to a well-developed system of heuristics that take into account both general and specific features of certain groups of interfaces.

Keywords: expert evaluation, heuristic evaluation, evaluation methods, user interface, expert's weighting factor.

АВТОМАТИЗИРОВАННАЯ ГЕНЕРАЦИЯ ВЕСОВЫХ КОЭФФИЦИЕНТОВ СПОСОБОВ ОТБОРА ПРИ ФОРМИРОВАНИИ ПУЛА РЕГРЕССИОННЫХ ТЕСТОВ

Нижегородский государственный технический университет им. Р.Е. Алексеева

Разработана методика автоматизированного выбора весовых коэффициентов отбора регрессионных тестов на основе анализа предыдущих проходов регрессионного тестирования. Разработанная методика интегрирована в программный комплекс по формированию пула регрессионных тестов.

Ключевые слова: регрессионное тестирование, автоматизированный отбор тестов, тестирование многокомпонентных информационных систем, методики отбора регрессионных тестов.

Введение

В настоящее время практически любая деятельность человека так или иначе использует различного рода информационные системы (ИС) [1]. Количество ИС непрерывно возрастает, также, как и количество сфер их применения, что влечет за собой необходимость в сокращении сроков разработки ИС, высокую нагрузку на организации, разрабатывающие ИС и значительный дефицит квалифицированных кадров, задействованных в процессе разработки и поддержки в течение жизненного цикла [2] ИС. Отдельно необходимо отметить, что при сокращении издержек на разработку и поддержку ИС компании-разработчики стремятся избежать снижения качества разрабатываемой ИС за счет применения современных методологий организации процессов разработки и тестирования [3]. Очевидно, для получения первых версий ИС, которые способны выполнять основные бизнес-требования, необходимо, прежде всего, сфокусироваться именно на разработке – написании кода программы. При этом, зачастую, в погоне за оптимизацией расходов и по причине некорректного планирования разработки, процессу тестирования ИС уделяется недостаточное количество времени несмотря на то, что выпуск на рынок некачественного ПО может привести к отсутствию интереса к ИС с точки зрения конечных пользователей и, следовательно, к неспособности окупить понесенные на разработку ИС затраты.

Процесс поиска ошибок в современных ИС требует от инженеров-тестировщиков или QA (quality assurance) глубоких знаний как в сфере именно «тестирования» [4], так и в смежных с ней сферах – разработки, аналитики, поддержки, маркетинга, понимания принципов передачи различного рода данных между системами [5] и так далее. При этом, в процессе формирования команды разработки ИС, необходимость в высокой квалификации именно инженеров по тестированию может быть неочевидна руководству проекта. По этой причине, при формальном наличии специалистов по проверке корректности ИС, разрабатываемый программный продукт рискует попасть на рынок без квалифицированной проверки на наличие в нем ошибок. Кроме того, даже при наличии в составе команды разработки инженеров-тестировщиков необходимой компетенции, возможна ситуация, когда из-за ошибок планирования для проведения всех необходимых проверок перед выпуском ИС в коммерческое использование, просто нет времени, что влечет за собой необходимость в первоочередном запуске самых важных и эффективных методик проверки качества ИС [6] и поиска ошибок [7].

В настоящее время существует большое количество способов поиска ошибок в ИС в течении всего жизненного цикла разработки ИС. При этом, с использованием в проекте по разработке ИС наиболее популярной в настоящее время Итеративной модели [8] (или её вариации – Спиральной модели [9], Scrum [10] и т.д.), наиболее важной с точки зрения

поиска ошибок становится Регрессионное тестирование (РТ) [11]. Запуск РТ очень важен с точки зрения обнаружения ошибок в областях кода, которые не подвергались в явном виде изменению в рамках текущей итерации разработки ИС, но могли потерять корректность функционала из-за изменений, внесенных в смежные с ними блоки кода. Это особенно актуально при разработке многомодульных ИС [12] и мобильных приложений [13].

Для повышения эффективности РТ, а также снижения издержек на отбор сценариев для РТ, необходимо разработать методику формирования пула РТ, которая учитывала бы результаты предыдущих запусков регрессионного тестирования.

Методика автоматизированной генерации весовых коэффициентов способов отбора при формировании пула регрессионных тестов

Одной из наиболее распространенных в настоящее время подхода разработки ИС является CI/CD–Continues integration/Continues delivery – Непрерывная интеграция/Непрерывная поставка [14]. Этот подход подразумевает создание новой версии ИС или ее составной части при внесении любого изменения или исправления. Особенно успешно этот подход применяется при использовании итеративной модели разработки для весьма популярных в настоящее время многомодульных ИС, однако использование этого подхода предъявляет крайне высокие требования к организации процедур тестирования. Особое внимание в процессе проверки качества ИС с многомодульной архитектурой требуется уделять РТ в связи с неявным влиянием на смежные модули внесенных в код изменений. При этом, для решения основных задач формирования пула РТ – Отбор, Минимизация, Приоритезация [15]– от инженеров-тестируемых требуются глубокие знания архитектуры проверяемой ИС, хорошее знание методик отбора тестов для РТ и перечня тестов, пригодных для запуска в рамках РТ.

Наиболее надежный способ поиска ошибок в рамках РТ является запуск всех доступных регрессионных тестов, но, к сожалению, в стандартной для современных проектов по разработке ИС ситуации, когда времени на проведение РТ и тестовых инженеров недостаточно, этот способ неприменим. Более того, при использовании подхода CI/CDвыполнение всех РТ может не иметь смысла, если таковое выполнение занимает значительное время –в процессе постоянного внесения изменений, выполненные некоторое время назад регрессионные тесты не позволяют сделать вывод о наличии ошибок в модулях многомодульной ИС уже после внесения одного-двух изменений в код ИС.

Задача формирования пула регрессионных тестов, учитывая все изменения, внесенные в программный код ИС, различные методики отбора РТ, особенности архитектуры ИС и т. д. может быть решена при помощи методики динамического отбора тестов [16]. Указанная методика позволяет формировать пул РТ, учитывающий изменения, внесенные в программный код ИС за произвольный период времени, применять различные методики отбора регрессионных тестов высокого уровня и ранжировать отобранные тесты в зависимости от значимости теста с точки зрения важности для принятия решения о пригодности ИС к коммерческому использованию.

Однако, данная методика предъявляет достаточно высокие требования к экспертной группе, которая определяет весовые коэффициенты для каждой из методик отбора РТ, что может быть недоступно в случае недостаточной квалификации инженеров, входящих в экспертную группу. Очевидно, что существует необходимость в разработке методики, которая бы позволяла применять методику динамического отбора тестов без глубоких знаний методик отбора регрессионных тестов и архитектуры разрабатываемой ИС. Для этого можно использовать методику анализа результатов предыдущих запусков РТ разрабатываемой ИС.

В процессе анализа предыдущих запусков РТ можно определить количество найденных ошибок по каждому тестовому сценарию (табл. 1).

Таблица 1

Оценка эффективности регрессионных тестов

Идентификатор теста	Количество обнаруженных ошибок
SQM-5435	3
SQM-2548	2
SQM-2854	2
SQM-4387	2
SQM-3984	1

Далее – соотнести тесты с методиками отбора, которыми эти тесты были отобраны и оценить таким образом количество обнаруженных ошибок по каждой методике отбора тестов РТ (табл. 2). При этом необходимо учесть, что один и тот же тест может быть отобран несколькими методиками отбора.

Таблица 2

Оценка эффективности методик отбора

Методика отбора	Количество обнаруженных ошибок
Тесты компонентов первого уровня влияния	12
Неуспешные тесты предыдущего РТ	7
Тесты высокого приоритета	6
Тесты, которые не запускались более 60 дней	3
Тесты на модуль отчетов	1

Таким образом, можно вычислить весовой коэффициент (ВК) методики отбора регрессионных тестов предыдущего РТ. Для вычисления итоговой значимости каждого теста ВК может быть использован как в абсолютных цифрах, так и приведен к некоему общему диапазону – от нуля до единицы, по формуле (1).

$$W_m = \frac{B_m}{B_1 + B_2 + \dots + B_n} \quad (1)$$

где

W_m – ВК методики отбора m ;

B_1 – количество найденных ошибок в тестах, отобранных методикой 1;

B_2 – количество найденных ошибок в тестах, отобранных методикой 2;

B_n – количество найденных ошибок в тестах, отобранных методикой n ;

n – количество методик отбора, которые позволили обнаружить ошибки в ИС.

В длительных проектах по разработке ИС, в которых применяется итерационный подход к разработке, может накопиться значительный объем данных по запускам РТ, которые также могут быть использованы при вычислении ВК методики отбора. Это может быть сделано простым суммированием значимостей методик отбора за определенное количество предыдущих запусков РТ (2)

$$W_a = \sum(W_1 + W_2 + \dots + W_n) \quad (2)$$

где

W_a – суммарный ВК методики отбора РТ;

W_1 – ВК методики отбора, найденный формулой (1) в запуске РТ 1;

W_2 – ВК методики отбора, найденный формулой (1) в запуске РТ 2;

W_n – ВК методики отбора, найденный формулой (1) в запуске РТ n ;

n – количество запусков РТ, которые были подвергнуты анализу.

При этом очевидно, что актуальность ВК методики отбора, вычисленная по результатам новых РТ выше, чем у более старых, что также может быть учтено с помощью коэффициентов актуальности анализа, принятых в диапазоне от нуля до единицы (3).

$$W_a = \sum(k_1 W_1 + k_2 W_2 + \dots + k_n W_n) \quad (3)$$

где

W_a – суммарный ВК методики отбора РТ;

k_1 – коэффициент методики 1;

W_1 – ВК методики отбора, найденный формулой (2) в запуске РТ 1;

k_2 – коэффициент методики 2;

W_2 – ВК методики отбора, найденный формулой (2) в запуске РТ 2;

k_n – коэффициент методики n;

W_n – ВК методики отбора, найденный формулой (2) в запуске РТ n;

n – количество запусков РТ, которые были подвергнуты анализу.

Количество запусков РТ, которое необходимо анализировать специфично для каждого из проектов и зависит от множества факторов, например, частоты запусков РТ, величины функционального прироста между запусками РТ, количества других методик тестирования ИС, которые применяются в процессе разработки ИС и так далее.

Необходимо учесть, что анализ предыдущих запусков РТ никак не учитывает специфику актуальной версии ИС, на которой планируется запуск РТ (какие модули были обновлены, какой новый функционал был внедрен, какие ошибки исправлены и т. д.), поэтому необходимо учитывать и мнение экспертной группы, которая принимает консолидированное решение о значении ВК для каждой из методик отбора РТ для текущего релиза. Результирующее значение ВК каждой из методик отбора может быть вычислено как сумма весового коэффициента предыдущих запусков РТ по формуле (3) и весового коэффициента, определенная экспертной группой (4).

$$W_t = \sum(w_i + W_i) \quad (4)$$

где

W_t – итоговый ВК методики отбора РТ;

w_i – ВК методики отбора, найденный формулой (3);

W_i – ВК методики отбора, определенный экспертной группой.

Применение методики автоматизированной генерации весовых коэффициентов способов отбора при формировании пула регрессионных тестов

Приведенная методика была имплементирована в программный комплекс ранжирования регрессионных тестов [17], который способен взаимодействовать с системами управления тестированием [18] посредством API [19] (рис. 1).

Key	Priority	Labels	Type	Executor	Execution date	Status	Defects	Last defect name	Weight
SBJ-37	Medium	[CICD]	Test	Jenkins CI Pip	03.15.19	Passed	16	Stable	83.3
SBJ-22	Medium	[CICD]	Test	Jenkins CI Pip	03.15.19	Passed	11	New functionality	75.1
SBJ-11	Highest	[CCN]	Test	Jenkins CCN Serv	03.15.19	Running	5	Stress	61.9
SBJ-21	Low	[MANUAL]	Test	Menister Alfou	03.15.19	Fail	7	Automatic	55.2
SBJ-16	High	[DAYLY]	Test	DailyBuild Server ...	03.16.19	Stoped	3	Stable	36.9
SBJ-19	Medium	[CICD]	Test	Jenkins CI Pip	03.16.19	Passed	3	Stress	35.3
SBJ-20	Low	[MANUAL]	Test	Menister Alfou	03.16.19	Passed	5	Manual	21.1
SBJ-35	Medium	[CICD]	Test	Jenkins CI Pip	03.16.19	Passed	6	Automatic	20.3
SBJ-33	Hign	[DAYLY] SBJ-03.16D	Test	DailyBuild Server ...	03.16.19	Passed	6	Automatic	16.1
SBJ-34	Medium	[CICD] SBJ-31	Test	Jenkins CI Pip	03.16.19	Running	4	New functionality	9.6
SBJ-40	Medium	[CICD] SBJ-32	Test	Jenkins CI Pip	03.16.19	Passed	3	Stress	4.7
SBJ-25	Medium	[CICD] SBJ-35	Test	Jenkins CI Pip	03.16.19	Passed	2	Stress	3.7
SBJ-26	High	[DAYLY] SBJ-03.17D	Test	DailyBuild Server ...	03.17.19	Fail	2	New functionality	3.6
SBJ-17	Low	[MANUAL] SBJ-6M	Test	Andy Lurie	03.17.19	Running	2	New functionality	3.6
SBJ-18	Medium	[CICD] SBJ-38	Test	Jenkins CI Pip	03.17.19	Passed	3	Stable	3.5
SBJ-13	Medium	[CICD] SBJ-39	Test	Jenkins CI Pip	03.17.19	Passed	3	New functionality	2.99
SBJ-39	Medium	[CICD] SBJ-39	Test	Jenkins CI Pip	03.17.19	Fail	2	Automatic	2.1
SBJ-30	Low	[MANUAL] SBJ-9M	Test	Menister Alfou	03.17.19	Running	1	New functionality	1.1
SBJ-26	Medium	[CICD] SBJ-40	Test	Jenkins CI Pip	03.18.19	Passed	1	Manual	1
SBJ-27	High	[DAYLY] SBJ-03.18D	Test	DailyBuild Server ...	03.18.19	Fail	1	Automatic	1

Рис. 1. Интерфейс программного комплекса ранжирования регрессионных тестов

Программный комплекс позволяет выполнить настройки глубины анализа выполненных ранее запусков, коэффициенты значимости и так далее (рис. 2).

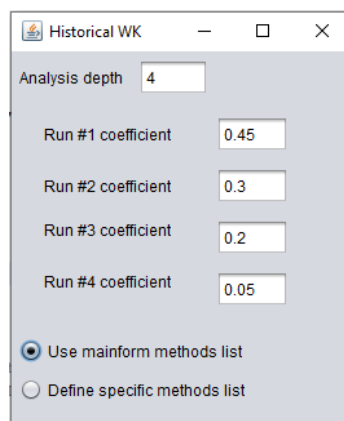


Рис. 2. Настройка автоматизированной генерации весовых коэффициентов

Программный комплекс с имплементированной методикой генерации ВК способов отбора был применен для формирования пула регрессионных тестов для двух запусков РТ при разработке ИС с высоким функциональным приростом и для трех запусков РТ при разработке ИС со средним функциональным приростом. В результате был получен прирост обнаруженных в рамках РТ ошибок на 4% для ИС с высоким функциональным приростом и около 2% для систем со средним функциональным приростом (рис. 3).

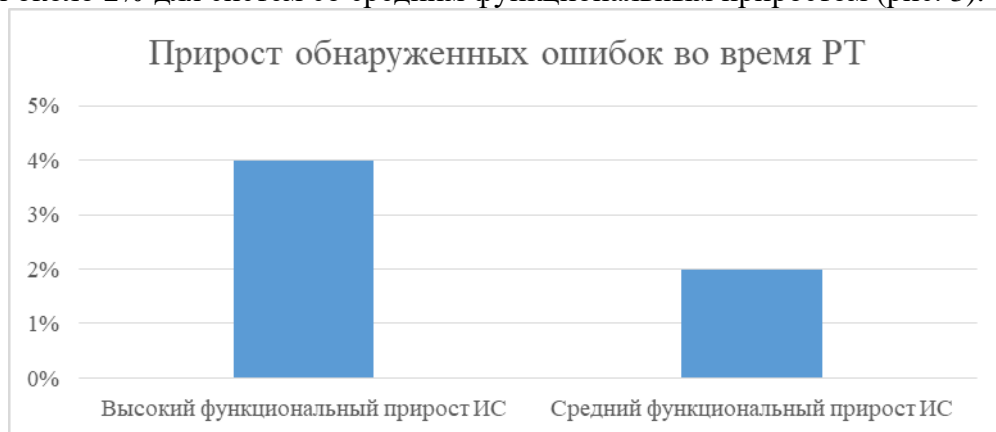


Рис. 3. Прирост обнаруженных ошибок во время РТ с автоматизированной генерацией ВК методик отбора

Выводы

Применение методики автоматизированной генерации ВК способов отбора для формирования пула регрессионных тестов на основании анализа результатов выполненных ранее регрессионных тестов позволяет помочь экспертной группе в принятии решения о значении ВК используемых методик отбора и повысить количество обнаруживаемых ошибок в процессе выполнения РТ. В дальнейших исследованиях необходимо разработать математическую модель вычисления ВК способов отбора тестов для формирования пула РТ, которая бы учитывала важность обнаруженных ошибок и потенциального влияния этих ошибок на смежные компоненты или блоки кода.

Библиографический список

1. Кузнецов С.Д. Большая Российская энциклопедия. – 2013 / [Электронный ресурс]. –Режим доступа: URL: https://bigenc.ru/technology_and_technique/text/3444940(дата обращения: 04.02.2021).
2. Национальный стандарт Российской Федерации ГОСТ Р 56136-2014 Управление жизненным циклом продукции военного назначения. Термины и определения. Переиздание 11.2016 с. 20 (с. 5-6).

3. McHugh, Martin; McCaffery, Fergal; Coady, Garret (4 November 2014). Mitasiunas, Antanas; Rout, Terry; O'Connor, Rory V.; et al. (eds.). *An Agile Implementation within a Medical Device Software Organization. Software Process Improvement and Capability Determination. Communications in Computer and Information Science.* 477. pp. 190–201. doi:10.1007/978-3-319-13036-1_17. ISBN 978-3-319-13035-4.
4. Гленфорд Майерс, Том Баджетт, Кори Сандлер. *Искусство тестирования программ.* Третье издание. 2019. с. 272. (с. 20-21) ISBN 978-5-907144-37-8.
5. Филинских А.Д., *Информационная метрика передачи и восстановления геометрических моделей в профессиональных программных средах,* диссертация кандидата технических наук, 05.13.17 - Теоретические основы информатики, защищена 26.12.13, утверждена 23.06.14 – Нижний Новгород, 2013 – 180 с.
6. Савин Р. *Тестирование Дот Ком или Пособие по жестокому обращению с багами в интернет-стартапах.* 2017. с. 312. ISBN 978-5-4485-4551-1.
7. SadhnaGoyal. *Influence of bugs on a software project.* KS OmniScriptumPublishin. 2010. с. 112. ISBN 9783639251746.
8. Larman C., Basili V.R. *Iterative and Incremental Development: A Brief History // Computer.* - 2003. – Vol.36, №6. – P.47-56.
9. Syed NaimatullahHussain and Ahmad TasnimSiddiqui. *Software Requirements and Process Models.* OmniScriptum Publishing KS. 2017. с. 68.
10. Сазерленд Джефф. *Scrum. Революционный метод управления проектами.* Манн, Иванов и Фербер. 2016. с. 288. ISBN 978-5-00146-509-6.
11. Yoo S. and Harman M. *Regression testing minimization, selection and prioritization: A survey.* King`sCollegeLondon — 2007. — с. 60.
12. Balalaie, A.; Heydarnoori, A.; Jamshidi, P. *Microservices Architecture Enables DevOps: Migration to a Cloud-Native Architecture (англ.) // IEEE Software: journal.* — 2016. — 1 May (vol. 33, no. 3). — P. 42—52. — ISSN 0740-7459. — DOI: 10.1109/MS.2016.64.
13. Гуляева У. И., Филинских А. Д. *Метрики разработки мобильного приложения.* Сборник материалов 28-й Всероссийской научно-практической конференции по графическим информационным технологиям и системам. Нижегородский государственный технический университет им. Р.Е. Алексеева. 2018. с. 134-138.
14. Pete Goodliffe. *chapter 22: The curious case of the frozen code // Becoming a Better Programmer: A Handbook for People Who Care About Code.* — "O'Reilly Media, Inc.", 2014-10-03. — С. 195 - 203. — 362 с. — ISBN 9781491905586.
15. Yoo S. and Harman M. *Regression testing minimization, selection and prioritization: A survey.* King`sCollegeLondon — 2007. — с. 60.
16. Zarubin, I.B., Filinskikh, A.D. *Method of dynamic selection of regression tests during developing multimodule information systems in conditions of CI/CD.* Conference Paper. CEUR Workshop Proceedings. 2020. 2763, стр. 226–231.
17. Зарубин И.Б., Долбунов Д.А., Филинских А.Д., Глумова Е.С., Гуляева У.И. *Система отбора регрессионных тестов. Свидетельство о государственной регистрации программы для ЭВМ RU 2020665635. Дата регистрации 24.11.2020.*
18. David Harned. *Hands-On Agile Software Development with JIRA.* Packt Publishing. 2018. с. 158. ISBN9781789532135.
19. Мехди Меджуи, Эрик Уайлд, Ронни Митра, Майк Амундсен. *Непрерывное развитие API. Правильные решения в изменчивом технологическом ландшафте.* Питер. 2020. с. 272. ISBN 978-5-4461-1232-6.

I.B. Zarubin, A D. Filinskikh

**AUTOMATED GENERATION OF WEIGHT COEFFICIENTS OF SELECTION
METHODS WHEN FORMING A POOL OF REGRESSION TESTS**

Nizhny Novgorod state technical university n.a. R.E. Alekseev

The method of automated selection of weight coefficients for the selection of regression tests based on the analysis of previous regression testing passes is developed. The developed methodology is integrated into the software package for forming a pool of regression tests.

Keywords: regression testing, automated test selection, testing of multicomponent information systems, regression test selection methods.

УДК: 681.327.8

Ф.Г. Хисамов, Е.В. Мельников

ПРОВЕРКА ИДЕНТИЧНОСТИ ПРОВЕРОЧНЫХ СООТНОШЕНИЙ $(k + m, k)$ КОДА МАКСИМАЛЬНОЙ ДЛИНЫ ПРОЦЕССУ СИНХРОНИЗАЦИИ ПСП ПО ЗАЧЕТНОМУ ОТРЕЗКУ

Кубанский институт информзащиты, г. Краснодар

Псевдослучайные последовательности (ПСП) находят широкое применение в специальных широкополосных системах связи различного назначения для защиты от организованных помех. В таких системах связи ПСП на передающей и приемных сторонах должны вырабатываться синхронно, поэтому перед началом передачи информации осуществляется синхронизация датчиков ПСП. Первоначальная синхронизация датчиков ПСП может осуществляться при отсутствии организованных помех, так как злоумышленнику не будет известен пеленг передатчика. Однако если происходит ложная синхронизация датчиков ПСП, которая обнаруживается по отсутствию свертки сигнала на приемной стороне, возникает необходимость в повторной синхронизации. Повторная синхронизация датчиков ПСП с высокой вероятностью будет происходить в условиях организованных помех, так как злоумышленнику становится известен пеленг передатчика. Исходя из этого, при проектировании специальных широкополосных систем связи возникает необходимость уменьшения вероятности ложной синхронизации датчиков ПСП. До настоящего времени вопрос оценки вероятности ложной синхронизации ПСП оставался открытым. Это было обусловлено тем, что сложно моделировать вероятность ложной синхронизации датчиков ПСП при неизвестных законах распределения ошибок, приводящих к таким последствиям. Поэтому вопросы математического моделирования вероятности ложной синхронизации ПСП являются актуальными при проектировании специальных систем широкополосной связи.

Ключевые слова: псевдослучайная последовательность, вероятность ложной синхронизации, коды максимальной длины, зачетный отрезок, организованные помехи, злоумышленник.

Введение

В данной работе доказывается теорема идентичности синхронизации датчиков ПСП по методу зачетного отрезка проверочным соотношениям $(k + m, k)$ кода максимальной длины, где k -длина регистра сдвига датчика ПСП, а m – длина порога переключения приемного датчика ПСП в автономный режим работы. Тогда вероятность ложной синхронизации датчика ПСП достаточно легко находится по известной формуле для вероятности необнаруженной ошибки $(k + m, k)$ кода максимальной длины.

Известно, что системы управления как гражданскими, так и военными объектами могут активно подавляться со стороны конкурирующих фирм или вероятного противника [1]. Для защиты систем управления от организованных помех применяют специальные широкополосные системы связи, которые являются инвариантными к организованным помехам. Это достигается псевдослучайной манипуляцией передающего и приемного сигналов по закону ПСП, вырабатываемых синхронно на обеих сторонах связи. Поэтому в

таких системах связи перед началом передачи информации осуществляется синхронизация датчиков ПСП. Здесь гораздо опаснее становятся ложные запуски приемного датчика, чем не запуски [2]. Это обусловлено тем, что для обнаружения ложного запуска приемного датчика ПСП потребуется больше времени чем для фиксации его не запуска, что дает возможность злоумышленнику запеленговать передатчик и при повторной синхронизации поставить мощные помехи.

Действительно при принудительном запуске ложная синхронизация произойдет при приеме любого ЗОТ с необнаруженной ошибкой. При этом за счет разных заполнений датчики на передаче и приеме будут вырабатываться разные ПСП, что приведет через некоторое время к срабатыванию детектора качества модема [2] это обнаружится на приеме как "обрыв" канала.

Если это произойдет в однонаправленной системе, то часть ПСП, которая передавалась в течении времени от начала ложного пуска до срабатывания детектора модема, будет потеряна, а оставшейся части ПСП может оказаться недостаточной для вхождения в синхронизм повторно или ее не будет вообще, так как закончится передача ПСП.

Если ложный запуск произойдет в системах с решающей обратной связью, то с момента ложного пуска на противоположную сторону по обратному каналу после подтверждения будет передаваться ПСП, а по прямому каналу будет ожидать подтверждения о фазировании, которое не сможет выделиться из-за различных ПСП, подаваемых на модемы передачи и приема [3]. Следовательно, по обратному каналу на противоположную станцию начнут поступать "запросы". При этом система может воспринять отсутствие прямого канала как следствие воздействия организованных помех и перейти в специально предусмотренный для этой ситуации режим работы и только тогда обнаружится ложный запуск приемного датчика ПСП [1,3]. То есть и здесь процесс определения ложности пуска затягивается во времени, что также может позволить злоумышленнику создать оптимизированные помехи.

Аналогичная ситуация возникает и при автономном пуске датчика ПСП. Поэтому систему синхронизации датчиков ПСП, предназначенных для работы в условиях организованных помех, необходимо выбирать из условия уменьшения вероятности ложной синхронизации, при заданной вероятности правильной синхронизации.

Однако к настоящему времени нет приемлемых оценок для расчета ложной синхронизации ПСП. Это обусловлено тем, что в обычных системах связи, предназначенных для работы при отсутствии организованных помех, вероятность возникновения ложной синхронизации не оказывает практически никакого влияния на обеспечение надежной синхронизации ПСП. Поэтому расчет вероятности ложной синхронизации при проектировании таких систем не проводится, так как такой расчет связан с трудностями математического моделирования этого процесса из-за отсутствия законов распределения ошибок, приводящих к ложной синхронизации.

В данной работе путем доказательства идентичности синхронизации ПСП процессу декодирования $(k+m, k)$ -кода максимальной длины, вероятность ложной синхронизации датчика ПСП сводится к определению необнаруженной ошибки $(k+m, k)$ -кода максимальной длины. Это доказывается следующей теоремой.

Теорема. Процесс синхронизации по методу ЗОТ путем проверки на совпадение символов ПСП, поступающих из канала и от обратной связи приемного ЛРР датчика ПСП, эквивалентен выполнению проверочных соотношений $(k+m, k)$ кода максимальной длины.

Доказательство.

Синхронизацию ПСП по методу зачетного отрезка при m совпадений элементов, поступающих из канала связи и от обратной связи приемного ЛРР датчика ПСП можно описать следующим выражением:

$$\sum_{j=0}^k v_{i-j} h_{k-j} = 0, \quad i = k, k+1, \dots, k+m-1, \quad (1)$$

где: $v_i, i = 0, 1, 2, \dots, k, k+1, \dots, k+m-1$ - символы, полученные из канала связи; $h_k, k = 0, \dots, k$ - коэффициенты полинома $h(x)$, на котором реализованы передающий и приемный ЛРР датчиков ПСП, причем $h_0 = h_k = 1, h(x)$ - примитивный полином степени k .

Легко заметить [4,5], что соотношение (1) есть ни что иное, как равенство нулю синдрома линейного $(k+m, k)$ кода V с проверочной $m \times (k+m)$ матрицей H :

$$H = \begin{matrix} \left[\begin{array}{cccccccc} h_0 & h_1 & \dots & h_k & 0 & \dots & 0 \\ 0 & h_0 & \dots & h_{k-1} & h_k & \dots & 0 \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots \\ 0 & \dots & h_0 & h_{k-1} & h_k & \dots & \dots \end{array} \right] \begin{matrix} \uparrow \\ m \\ \downarrow \end{matrix} \end{matrix} \quad (2)$$

Действительно, легко убедиться, что (1) эквивалентно выражению:

$$V \cdot H^T = 0, \quad (3)$$

где: H^T - транспонированная матрица H из (2); $V = (V_0, V_1, \dots, V_{k+m-1})$ - кодовые слова $(k+m, k)$ кода.

Принимаемые символы $V_i, i = 0, \dots, k+m-1$, связаны с передаваемыми символами $b_i, i = 0, 1, \dots, k+m-1$, следующим очевидным соотношением [5]:

$$V_i = b_i \oplus e_i, \quad i = 0, 1, \dots, k+m-1, \quad (4)$$

где $e_i = 0$ или 1 - последовательность образцов ошибок в канале связи.

Следовательно, процесс синхронизации ПСП по методу зачетного отрезка полностью изоморфен декодированию $(k+m, k)$ -кода максимальной длины, что и требовалось доказать.

Заключение

На основании доказанной теоремы оценку вероятность ложной синхронизации ПСП свелась к известной оценке необнаруженной ошибки $(k+m, k)$ -кода максимальной длины. То есть, используя теоретический аппарат теории кодирования, удалось получить достаточно простое математическое выражение для расчета вероятности ложной синхронизации ПСП через вероятность необнаруженной ошибки $(k+m, k)$ кода максимальной длины, которая приведена ниже:

$$P_n = P_l = \frac{P_{no}}{P_{oo}} = \frac{\sum_{i=k}^{2k} N_i P^i (1-P)^{k-i}}{(1-P)^{2k}} = \frac{2^t}{(1-P)^{2k}}$$

где P - вероятность ошибки символа в канале; N_i - число кодовых слов веса i в «усеченной» M -последовательности (спектр кода); d_m - минимальное Хэмминговое расстояние кода; $t = \left\lceil \frac{d_m - 1}{2} \right\rceil$.

Библиографический список

1. Куприянов А.И. Радиоэлектронная борьба / А.И. Куприянов. - 4-е изд. - М.: Вузовская книга. 2016. - 360с.: ил.
2. Мельников Е.В. Оценка оптимального порога синхронного запуска приемного датчика АПСП коммерческих широкополосных систем связи. / ИНФОРМАТИКА и безопасность современного общества / научно-практический журнал № 2(8). - Краснодар: КИИЗ, 2019. - С. 11-13.

3. Мельников Е.В., Ажмухамедов И.М. Оценка вероятности ложной синхронизации АПСП широкополосных систем связи. / III Научный форум телекоммуникации: Теория и технологии ТТТ-2019. Проблемы техники и технологий телекоммуникаций ПТиТТ-2019: материалы XXI Международной научно-технической конференции. Казань, 18 – 22 ноября 2019 года. – Казань: КНИТУ-КАИ, 2019 – Т. 1 – 696 с. С. 59-60.
4. Коржик В.И., Финк Л.М. Помехоустойчивое кодирование дискретных сообщений в каналах со случайной структурой - М.: Связь, 1975.
5. Морелос-Сарагоса Р Искусство помехоустойчивого кодирования. Методы, алгоритмы, применение – М: Издательство: Техносфера, 2006. - 320 с

F.G. Khisamov, E.V. Melnikov

CHECKING THE IDENTITY OF THE MAXIMUM LENGTH CODE CHECK PROCESS FOR THE PSEUDO-RANDOM SEQUENCE SYNCHRONIZATION PROCESS BY CREDIT

The Cuban Institute of Information Protection

Purpose: This article is dedicated to assessing the probability of false synchronization of pseudo- random sequence in broad band systems, using the theory of interference-resistant coding.

Design/methodology/approach: Methodology consists in strict proof of the identity of synchronization of pseudo- random sequence by the method of credit segment decoding the code of maximum length. The probability of false synchronization of pseudo- random sequence is reduced to a known estimate of an undetected error $(k-m, k)$ -the code of maximum length.

Findings: When designing special broadband communication systems, it is necessary to reduce the probability of false synchronization of pseudo-random sequence sensors. Based on the proven identity theorem of the synchronization of pseudo-random sequence by the method of credit segment decoding the code of maximum length, it was possible to obtain a fairly simple mathematical expression to calculate the probability of false synchronization of pseudo-random sequence.

Originality/value: Until now, the question of assessing the probability of false synchronization of pseudo- random sequence has remained open. This was due to the fact that it is difficult to simulate the probability of false synchronization of pseudo- random sequence sensors under unknown laws of distribution of errors leading to such consequences. Therefore, the estimates of the probability of false synchronization of pseudo- random sequence are relevant in the design of special broadband systems of various purposes.

Keywords: pseudo-random sequence, probability of false synchronization, maximum length codes, credit segment, organized interference, attacker.

АНАЛИЗ ЖИВУЧЕСТИ СИСТЕМЫ УПРАВЛЕНИЯ СЕРТИФИКАТАМИ КЛЮЧЕЙ ПРОВЕРКИ ЭЛЕКТРОННОЙ ПОДПИСИ НА ОСНОВЕ РАЗЛИЧНЫХ АРХИТЕКТУР

Краснодарское высшее военное училище имени генерала армии С.М. Штеменко

В статье проведён сравнительный анализ архитектур, реализующих механизмы обеспечения живучести применительно к системе управления сертификатами ключей проверки электронной подписи (далее СУ СКП ЭП). Обосновано, что наиболее эффективной реализацией резервирования по показателю ресурсозатратность обладает децентрализованная архитектура, используемая в технологии распределённых реестров.

Ключевые слова: сравнительный анализ, механизмы обеспечения живучести, одноранговая архитектура, распределённая система удостоверяющих центров, ДНТ.

Введение

Для того чтобы СУ СКП ЭП обладала свойством живучести, она должна реализовывать, по крайней мере, один из механизмов, обеспечивающих ее реализацию. Следовательно, к «живучим» СУ СКП ЭП будем относить такие системы, в которых наиболее эффективно реализованы механизмы восстановления или (и) адаптации на основе резервирования [1], [2], [3].

В настоящее время к подобным системам можно отнести:

- 1) системы, реализующие общее резервирование (системы со статической архитектурой, системы с динамически перестраиваемой архитектурой);
- 2) Системы с поэлементным резервированием – высоконадежные серверы.

Проведем анализ архитектур с целью выявления архитектуры, которая наиболее эффективно реализует резервирование. При выборе архитектуры основополагающим показателем будет являться ресурсозатратность.

Анализ живучести СУ СКП ЭП на основе статической архитектуры

В настоящее время СУ СКП ЭП строится на основе статической архитектуры и реализует только один механизм обеспечения живучести - восстановление. Если рассмотреть программно-технический комплекс каждого отдельного удостоверяющего центра (далее ПТК УЦ), то он чаще всего использует две модели резервирования:

- с включённым горячим резервом;
- на основе периодического резервного копирования данных.

В первой модели резервирования отсутствует единая точка отказа, приводящая к выходу из строя всю систему. При отказе одного из серверов обработки данных его нагрузку берут на себя оставшиеся сервера. Пользователи могут получать доступ к службам ПТК УЦ независимо от количества оставшихся работоспособных серверов обработки данных, что обеспечивается на основе виртуализации сетевого адреса ПТК УЦ (т.е. все сервера обработки данных имеют единый сетевой адрес, возможные конфликты обработки запросов решаются путём применения выделенной локальной сети).

Главным недостатком рассматриваемой модели является то, что она имеет ограниченную способность по компенсации отказов, которая зависит от количества серверов, находящихся в горячем резерве у ПТК УЦ. Если ПТК УЦ содержит два сервера, то, соответственно, если они оба выйдут из строя, то он перестанет функционировать. Выход из строя хотя бы одного ПТК УЦ в виду построения иерархической модели доверия приведет СУ СКП ЭП к отказу в обслуживании.

Следующий недостаток статической архитектуры состоит в том, что ее использование предполагает «отчуждение» под нужды ПТК УЦ определенного

количества серверов. То есть, если ПТК УЦ бездействует или функционирует с минимальной загрузкой вычислительных ресурсов, то это означает, что незадействованные серверы также простаивают, что негативно сказывается на ресурсозатратности. Кроме того, к недостаткам рассматриваемой архитектуры можно отнести необходимость покупки дополнительного дорогостоящего оборудования.

Недостатком второй модели является тот факт, что информацию, записанную на сервер в текущий момент времени при реализации деструктивных воздействий (например, при успешном функционировании вредоносного программного обеспечения), невозможно будет восстановить. Кроме того, на время осуществления резервного копирования СУ СКП ЭП будет недоступна для пользователей.

Анализ эффективности резервирования СУ СКП ЭП на основе технологий с динамически перестраиваемой архитектурой

Системы с динамически перестраиваемой архитектурой можно подразделить на три класса:

1. Централизованные.
2. Децентрализованные.
3. Гибридные.

Проанализируем эффективность резервирования данных классов на примере:

1. Облачной архитектуры.
2. Архитектуры на основе туманных вычислений.
3. Централизованной Grid-архитектуры.
4. Децентрализованной Grid-архитектуры.
5. Децентрализованной архитектуры распределенного реестра.

Анализ живучести РСОД на основе архитектуры облачных вычислений

Облачная архитектура является представителем 1 класса (рис. 1). Облачная архитектура обладает такими механизмами обеспечения живучести как адаптация и восстановление, при этом является гибридной моделью, реализующей методы структурного, функционального и информационного резервирования.

Большинство работ [4] направлены на исследование такого свойства данной архитектуры как производительность, так как конечные пользователи для своих нужд используют только вычислительные мощности облака. Обычно крупные облачные центры обработки данных (далее ЦОД) состоят из тысяч серверов обработки запросов, которые образуют виртуальную среду благодаря высокоскоростным линиям связи и специальному программному обеспечению, где конечному пользователю нет необходимости задумываться, к какому физическому серверу или множеству серверов он подключён.

Не смотря на все преимущества данной архитектуры, у неё есть и свои недостатки:

1. В виду того, что все заявки от всех пользователей всегда проходят через маршрутизатор или множество маршрутизаторов ЦОД, данные устройства и линии связи, по которым пользователи устанавливают соединение с ними, являются единой точкой отказа.

2. Развёртывание инфраструктуры облака является очень дорогостоящим и оправдывает себя только в случае необходимости огромной вычислительной мощности для ресурсозатратных приложений.

3. Облачная архитектура не функционирует в реальном времени, а при реализации деструктивных воздействий на узлы и каналы связи временная задержка может не соответствовать требованиям по оперативности обработки запросов, предъявляемых к СУ СКП ЭП.

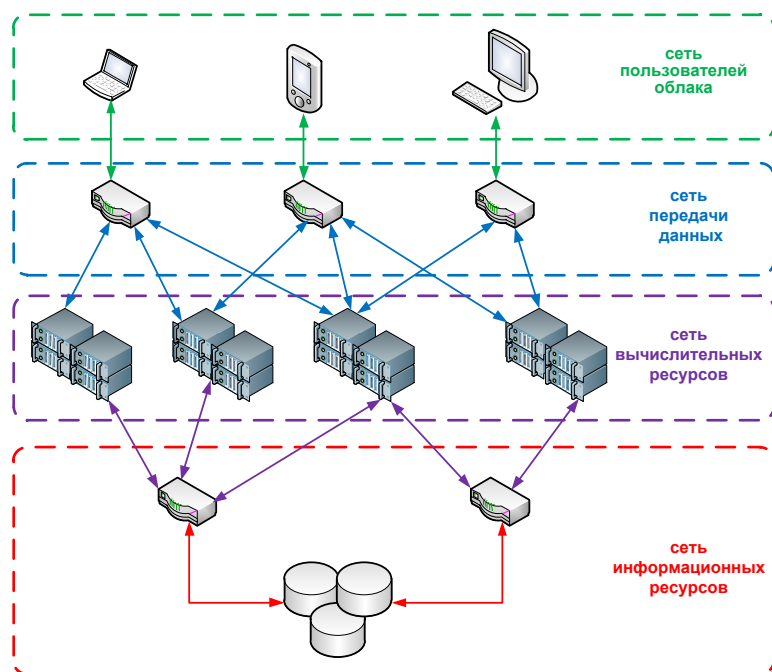


Рис. 1. Облачная архитектура

Анализ эффективности резервирования СУ СКП ЭП на основе архитектуры туманных вычислений

Архитектура туманных вычислений не создавалась в противовес облачным вычислениям, а скорее дополняла данную архитектуру с целью уменьшения временной задержки обработки запросов от конечных пользователей на основе установки в непосредственной от них географической близости соответствующих туманных узлов, благодаря чему позволила перераспределить вычислительные мощности облака и снизить нагрузку на каналы связи, используемые для соединения с ним. Тем не менее, построение СУ СКП ЭП на основе данной архитектуры на данный момент является ресурсозатратным в виду того, что она остаётся централизованной иерархической системой, осуществляющей структурное резервирование, что делает её ещё более дорогостоящей для развёртывания в сравнении с облачной архитектурой.

Анализ эффективности резервирования СУ СКП ЭП на основе централизованной Grid-архитектуры

Программное обеспечение Grid – OGSA (Open Grid Service Architecture) базируется на объектно-ориентированной модели [5,6]. В рассматриваемой архитектуре в качестве основного объекта Grid рассматривают службу. Посредством удалённого обращения к методам службы приложение получает определённый вид обслуживания. Таким образом, унифицируются различные функции доступа к вычислительным ресурсам, ресурсам хранения, базам данных и к любой их программной обработке.

Grid-архитектура создается на основе динамического объединения серверов и позволяет обеспечить доступ пользователей к их разделяемым вычислительным и информационным ресурсам. Данная архитектура реализует механизм адаптации системы на основе компенсации отказов каких-либо серверов в обработке запросов пользователей. Так, в случае отказа одного из них система может перенаправить пользователя на другой сервер, выполняющий аналогичные функции.

Частной реализацией функционирования Grid архитектуры является торрент-система: раздающие узлы выступают в роли серверов, исполняющих функцию распространения информационных ресурсов. Обращаясь к определённой службе (торрент-трекер), узлы-получатели запрашивают данные от Grid-диспетчера о местонахождении определённых информационных ресурсов (ИР), которые разбиты на сегменты (рис. 2).

Данные информационные системы служат для наиболее быстрого доступа к получению данных большого объёма. Торрент-трекеры хранят ссылки на информационные ресурсы соответствующих пользователей в виде хэш-кодов каждого сегмента, выступая доверенной стороной, обеспечивающей целостность запрашиваемого ресурса. Тем не менее, если в системе отсутствуют раздающие узлы, имеющие один и тот же информационный ресурс, то узлы-получатели не смогут воспользоваться услугами нужной ему службы. Таким образом, данная архитектура осуществляет только мониторинг имеющихся вычислительных или информационных ресурсов, вследствие чего система может работать нестабильно, таким образом, имея более длительную временную задержку, чем в облачной архитектуре.



Рис.2. Централизованная GRID-архитектура

Анализ эффективности резервирования СУ СКП ЭП на основе децентрализованной Grid-архитектуры

В отличие централизованной Grid-архитектуры в протоколе «Kademlia» [7] торрент-трекеры отсутствуют (их функции выполняет каждый узел в сети на основе ведения единой децентрализованной хэш-таблицы (DHT)). Оценивая эффективность резервирования данной архитектуры, можно выделить следующие преимущества:

- является наименее ресурсозатратной, так как нет необходимости в создании инфраструктуры серверов;
- отказы узлов или их временный выход из системы не приведёт её в неработоспособное состояние;
- является наиболее оперативной, так используется максимальная общая пропускная способность системы каналов связи.

В то же время применительно к СУ СКП ЭП существует и недостаток данной архитектуры. Отсутствие алгоритмов консенсуса (конечной согласованности) между узлами сети о записи данных в DHT соответствующими доверенными узлами-

регистраторами приведёт к риску осуществления несанкционированного доступа к системе, тем самым нарушая целостность записываемых в неё данных.

Анализ эффективности резервирования СУ СКП ЭП на основе технологии распределённых реестров

В настоящее время известна технология распределённого хранения и обмена данными без возможности их подмены или удаления, именуемая технологией распределённых реестров (блокчейн). Технология блокчейн является частным случаем технологии распределённых реестров и применяется только в криптовалютных системах. Структура записи данных имеет вид цепной блочной записи транзакций с использованием таких криптографических преобразований как функция хеширования и электронная подпись. Функционирование автоматизированных информационных систем на основе данной технологии осуществляется децентрализованным способом на основе исполнения алгоритма достижения консенсуса между её узлами. Блокчейн-система легко адаптируется к выходу из строя или целенаправленному отключению её узлов в результате разрыва соединений с ней в случайные моменты времени. Процесс реконфигурации взаимодействующих работоспособных узлов на основе установления новых соединений в виду распределённого доверия проходит в автоматическом режиме без использования единого центра управления, что значительно повышает живучесть системы по сравнению с централизованными архитектурами. При этом для обеспечения конечной согласованности о записи данных в распределённый реестр требуемая связность должна быть обеспечена не менее двумя третьими установленными соединениями между узлами от максимально возможного числа соединений. Вместе с этим функционирование данной системы невозможно без доверенных узлов-регистраторов, обеспечивающих контроль формирования структуры реестра и корректность содержащихся в нём транзакций. В случае уничтожения всех узлов данного типа или отсутствия с ними связи произойдёт отказ системы, пока не произойдёт выбор и назначение новых узлов (восстановление системы).

Заключение

Таким образом, проведённый анализ технологий на основе различных архитектур, изображенный на рис. 3 показал, что архитектура на основе технологии распределённых реестров для СУ СКП ЭП обеспечивает максимальную эффективность резервирования, а соответственно и ее живучесть. Следовательно, функционирование СУ СКП ЭП на основе выбранной архитектуры является наиболее предпочтительным вариантом и требует более детального анализа механизмов её функционирования.



Рис. 3. Анализ технологий с динамически перестраиваемой архитектурой для реализации резервной СУ СКП ЭП

Библиографический список

1. Мамиконов А.Г. Проектирование АСУ. Москва: Высшая школа, 1987.
2. Максименков А.В. Основы проектирования информационно-вычислительных систем и сетей ЭВМ. Москва: Радио и связь, 1991.
3. Малашенко Ю. Е., Рогожин В. С, Ферапонтов Е. В. Живучесть сетевых систем. Москва: ВЦ АН СССР, 1989.
4. Комухаев С. Тестирование производительности облачных инфраструктур [Электронный ресурс] URL: <https://www.sim-networks.com/blog/performance-of-cloud-infrastructure> (дата обращения 14.01.2021).
5. Ян Фостер, Карл Кессельман, Джеффри Ник. Grid-службы для интеграции распределенных систем. [Электронный ресурс] URL: <https://www.osp.ru/os/2003/01/182393> (дата обращения 14.01.2021).
6. Мамин Е. Kademia DHT: Основы [Электронный ресурс] [2010]. URL: <https://habr.com/ru/post/107342/> (дата обращения 14.01.2021).

A. S. Toroev, A. B. Sizonenko, I.P. Pavlov

ANALYSIS OF THE SURVIVABILITY OF THE ELECTRONIC SIGNATURE VERIFICATION KEY CERTIFICATE MANAGEMENT SYSTEM BASED ON DIFFERENT ARCHITECTURES

Krasnodar higher military school n.a. general S.M. Shtemenko

The article presents a comparative analysis of architectures that implement mechanisms for ensuring survivability in relation to the electronic signature verification key certificate management system (ES VKC MS). It is proved that the most effective implementation of reservation in terms of resource consumption has a decentralized architecture used in distributed ledger technology.

Keywords: comparative analysis, survivability mechanisms, peer-to-peer architecture, distributed system of certification centers, DHT.

ВИДЫ И ОПРЕДЕЛЕНИЯ КРИПТОГРАФИЧЕСКИХ СИСТЕМ

¹ Саровский физико-технический институт – филиал НИЯУ МИФИ,² РФЯЦ-ВНИИЭФ

В статье рассмотрены виды систем и различные варианты определения систем, в частности, криптографических, в рамках понятийного аппарата общей теории систем. Также предлагается дополненный и уточненный жизненный цикл данных систем.

Ключевые слова: система, системный подход, теория систем, виды систем, жизненный цикл систем.

Введение

Потребность в использовании понятия «система» возникала для объектов различной физической природы с древних времён: ещё Аристотель обратил внимание на то, что целое несводимо к сумме частей, его образующих. Термин «система» и связанные с ним понятия комплексного, системного подхода исследовались и подвергались осмыслению философами, биологами, психологами, кибернетиками, физиками, инженерами различных специальностей. Интерес к системным представлениям проявлялся не только как к удобному обобщающему понятию, но и как к средству постановки задач с большой неопределённостью [1].

Очевидно, что научный аппарат для анализа и изучения сложных систем призвана дать такая дисциплина, как системный анализ, сформировавшаяся в результате возникновения необходимости исследовать и проектировать сложные системы, управлять ими в условиях неполноты информации, ограниченности ресурсов и дефицита времени [2].

Рассмотрим несколько вариантов определений систем, в частности, криптографических систем, а также жизненный цикл данных систем.

Определение системы через термины теории множеств

Существует несколько десятков определений понятия система, которые приведены в работах [3-5]. Хотя в переводе с греческого языка слово система означает *состав*, система - это не просто множество элементов. Компоненты системы находятся во взаимосвязи, которые определенным образом организованы, благодаря чему обеспечивают целостность системы. Учитывая, что криптографическая система определяется как некоторое отображение множества сообщений на множество криптограмм, рассмотрим вариант определения системы через термины теории множеств.

При определении системы через термины теории множеств ее совокупности взаимосвязанных элементов соответствует следующее выражение

$$S \stackrel{\text{def}}{=} \langle V, R \rangle, \quad (1)$$

где V - множество элементов системы $V = \{V_1, V_2, \dots, V_n\}$; R - множество отношений между элементами системы $R = \{R_1, R_2, \dots, R_n\}$.

Так как система состоит из элементов различной природы, то систему можно задавать в виде

$$S \subseteq V_1 \times V_2 \times \dots \times V_n \text{ или } S \stackrel{\text{def}}{=} \langle V_1, V_1, \dots, V_n, R \rangle, \quad (2)$$

где V_i – множество однородных элементов некоторого вида, появляющихся при целевом разделении базовых элементов на их составляющие.

Иногда элементы системы разделяются на два типа множеств: множество входных объектов X , воздействующих на систему, и множество выходных результатов Y . В этом случае систему определяют как $S \subseteq X \times Y$.

Можно задать систему как функцию выходных объектов от входов $S: X \rightarrow Y$. Определению системы, в которое дополнительно введено множество свойств Q , характеризующих как отдельные элементы, так и всю систему в целом, соответствует запись

$$S \stackrel{\text{def}}{=} \langle V_1, V_1, \dots, V_n, Q, R \rangle, Q = Q_s \cup Q_v, \quad (3)$$

где Q_s – множество свойств системы; Q_v – множество свойств элементов.

Конструктивное определение системы, включающего наблюдателя N или лицо анализирующее объект или процесс в виде системы (в данном случае это может быть злоумышленник) и цель (совокупность или структуру целей) Z , может быть представлено как

$$S \stackrel{\text{def}}{=} \langle V_1, V_1, \dots, V_n, Q, R, N, Z \rangle. \quad (4)$$

Дополнительно в формулу могут быть введены: окружающая среда системы Sr (учитывающая деятельность злоумышленника и естественные и искусственные помехи в канале связи) и интервал времени ΔT , т.е. период, в рамках которого будет существовать система и ее цели:

$$S \stackrel{\text{def}}{=} \langle V_1, V_1, \dots, V_n, Q, R, N, Z, Sr, \Delta T \rangle. \quad (5)$$

Все это делает актуальным построение не только функциональных, но и множественных моделей информационных и криптографических систем и их составных элементов, а также исследование систем через их свойства.

Определение системы через ее свойства при структурно-статическом, функционально-динамическом и информационном подходах

Рассмотрим определение системы через ее свойства. Итак, систему, после определения ее состава и взаимосвязи элементов внутри системы и с внешней средой, построения множественных моделей, можно определять также через ее свойства и свойства ее элементов, к которым относятся:

- 1) при структурно-статическом подходе: целостность, делимость, взаимосвязь, иерархичность, взаимодействие и коммуникативность;
- 2) при функционально-динамическом подходе: динамичность, историчность, устойчивость и целеустремленность;
- 3) при информационном подходе: количество информации, неопределенность и энтропия.

Свойства и закономерности систем приведены на рис.1. Наименования подходов уточнены, а свойства проанализированы и дополнены по сравнению с аналогичной классификацией, описанной в работах [3-5].



Рис. 1. Свойства и закономерности систем

При структурно-статическом подходе в соответствии со свойствами целостности и делимости система всегда рассматривается как совокупность элементов и подсистем с учетом их взаимосвязей и иерархичности. Предел делимости определяется целями рассмотрения системы, в частности это относится к процессам рассмотрения конкретных

криптографических функций, шифров и алгоритмов, подробно рассмотренных в работах [6,7,8]. В качестве подсистем при таком рассмотрении могут выступать относительно самостоятельные части системы, обладающие определенной целостностью и функциональностью. Сама система или алгоритм могут рассматриваться как подсистема более крупной системы. Для любой системы применим принцип иерархичности, который в числе первых выделил и исследовал Л. фон Берталанфи [9].

Система в целом и ее элементы обладают определенными свойствами. Свойства элементов, взятых вне системы, отличаются от их свойств в рамках системы, из чего следует, что система в целом может обладать и обладает свойствами, которыми не обладают в сумме ее элементы

$$Q_s \neq \sum_{i=1}^n q_i, \quad (6)$$

где Q_s - свойства системы, а q_i - свойства i -го элемента. Из выражения следует, что система в целом не является просто суммой свойств составляющих ее частей, но это не означает независимости Q_s от q_i . Наоборот, свойства системы зависят от свойств составляющих ее элементов

$$Q_s = f(q_i), i = \{1, \dots, n\}. \quad (7)$$

Несводимость свойств системы к сумме свойств ее элементов приводит к понятию эмерджентности, т.е. появлению у системы принципиально новых качеств, не имеющих у ее элементов при их объединении в систему. Принцип эмерджентности обеспечивается за счет выполнения принципа взаимодействия: единство обеспечивается взаимодействием частей. Взаимодействие — это установление связей между элементами, когда изменение одного из них влечет изменение другого. Безусловно, свойства системы в целом проявляются во взаимодействии с окружающей средой, т.е. реализуется через внешние связи, но сами эти свойства возникают благодаря внутренним связям, составляющим структуру системы.

Следует отметить, что для обеспечения целостности системы необходимо, чтобы выполнялось следующее соотношение

$$W_{V_i} > W_{S_j}, \quad (8)$$

где W_{V_i} - суммарная мощность внутренних связей, а W_{S_j} - суммарная мощность внешних связей. Если внутренние связи ослабевают, то система распадается.

Следует отметить, что функционирование любой информационной и криптографической системы происходит в некоторой окружающей внешней среде, определенным образом взаимодействующей с системой. Среда — это совокупность объектов, изменение свойств которых влияет на систему, а также объектов, чьи свойства меняются в результате поведения системы. Особенно это проявляется при построении и анализе систем на базе известных криптографических алгоритмов, обладающих определенной стойкостью. Это понимание является сутью коммуникативности и является важным для процессов информационно-технического управления системами. Особо отметим, что необходимо изучать и анализировать внешнюю среду, прогнозировать ситуацию с тем, чтобы вовремя и адекватно реагировать на ее изменения, что особенно важно при стремительном ускорении технического прогресса и резком возрастании вычислительных возможностей [5,10].

Отметим, что при функционально-динамическом подходе следует учитывать, что любая система не является застывшим, неизменным образованием. В результате внешних и внутренних взаимодействий система находится в динамике и может быть подвержена постоянным изменениям. Состояние системы — это множество одновременно существующих свойств системы, которое проявляется в изменениях ее параметров с течением времени:

$$P = F(Q, t), \quad (9)$$

где P — показатель состояния системы, Q - совокупность параметров, t - время, F — функциональная зависимость.

Примером функционально-динамического подхода к анализу сложных информационных систем и их составляющих элементов является работа «Аксиоматические основы функций подстановки ряда факториальных множеств и их характеристики» [10], вышедшая в 2019 г.

В свою очередь, при функционально-динамическом подходе, когда одно состояние сменяется другим, это трактуется как событие, т.е. переход в новое состояние. Функционирование системы во время своей жизнедеятельности представляет собой цепочку событий или процесс, состоящий из последовательно сменяющих друг друга состояний.

Состояние системы в целом и ее основных частей в каждый момент времени зависит от предыдущего состояния, а также от внутренних и внешних взаимодействий. Динамическая система может находиться либо в равновесном состоянии, либо в состоянии перехода в новое состояние. Понятие равновесия определяется, как способность системы в отсутствии внешних воздействий сохранять свое состояние сколько угодно долго.

Целенаправленность системы отражает динамику и последовательность изменения ее состояния. Цель – это желаемый результат деятельности системы, зависящий от времени. Цель может быть задана как желаемое состояние для ограниченных систем или желаемое направление изменения системы. Однако отметим, что для сложных открытых систем, какими являются информационные и криптографические системы, действующие в постоянно меняющихся условиях и при изменяющихся целях, этого может оказаться недостаточно. Для них, кроме этого, необходимо определять, как тенденции, так и перспективы развития. Для этого необходимо увеличивать роль управления системой для достижения поставленной цели. Система является управляемой, если каждому воздействию управления соответствует строго определенное состояние ее параметров, для нас такими параметрами являются сообщения и ключевая информация. В работах по теории систем и системному анализу часто приводится классическая кибернетическая схема функционирования системы управления, содержащая объект управления и управляющую подсистему, которая задействована через обратную связь [10]. Данной схеме в полной мере соответствует модель цифрового устройства с характеристическими матрицами, предложенная в работе [11] при рассмотрении вопросов теории поля и защиты информации и приведенная в Приложении 2.

Принцип управления через обратную связь является одним из важнейших принципов управления, что подтверждает тезисы о том, что управление сложной системой может осуществляться только при получении информации о результатах реализации предыдущих управляющих воздействий.

Пределные возможности управления сформулированы У.Р. Эшби [12] суть которых заключается в следующем: для того чтобы создать систему, способную справиться с решением проблемы, обладающей определенным разнообразием, нужно, чтобы система имела еще большее разнообразие, чем разнообразие решаемой проблемы, или была способна создать себе это разнообразие. Из этого закона делаем вывод, что невозможно создать простую систему управления или защиты для сложной системы, действующей в недетерминированной среде [5,13]. Этот тезис подтверждается и в работах по исследованию общетеоретической концепции и архитектурных принципов построения распределенных адаптивных систем с функцией защиты данных [14] и обеспечению подлинности и целостности информации при проведении научных экспериментов [15]. Данные работы утверждены начальником филиала Военной академии РВСН им. Петра Великого в качестве учебного пособия для курсантов вузов и специалистов по направлению «Информационные системы и технологии».

Жизненный цикл системы

В процессе своей жизнедеятельности системы проходят определенные этапы, которые называются ее жизненным циклом. Этапы жизненного цикла, такие как создание, развитие, эксплуатация и упадок следуют в строгой последовательности и характеризуются определенными предсказуемыми состояниями. Отметим, что учет этого факта помогает правильно выбрать стратегию управления.

Например, если на этапе упадка жизненного цикла системы вовремя произвести анализ причин упадка, чем может являться устаревание элементов системы или изменение внешних условий функционирования системы, возрастание вычислительных возможностей злоумышленника, то система может получить новое развитие и можно будет продлить ее эксплуатацию, иначе придется проводить ее утилизацию, а затем новую разработку. С учетом этого предлагается уточненный и дополненный по сравнению с [5] вариант жизненного цикла системы, приведенный на рис. 2.

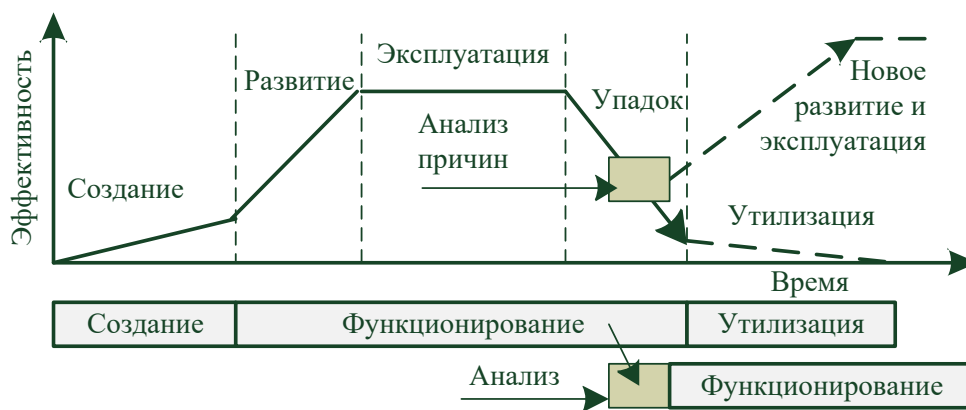


Рис. 2. Уточненный вариант жизненного цикла системы

Анализ причин может производиться и на этапе эксплуатации. Динамическая система в общем случае может находиться либо в равновесном состоянии, либо в состоянии перехода из состояния в состояние. Эти процессы для теории систем подробно описаны в работах [3-5]. Постоянно развивающиеся системы, к которым относятся информационные и криптографические системы, имеют некоторый предел, определяемый их структурой и свойствами. Основой этих характеристик является теоретическая и практическая стойкость криптографических систем [6,7].

Прохождение системами определенных стадий развития называется закономерностью историчности. Знание данной закономерности помогает правильно выбрать стратегию управления и развития. Исторические аспекты развития криптографических и информационных систем на ранних этапах нашли отражение в работах [17,18]. Один из основных выводов заключается в том, что основными функциями криптографических преобразований в них являются функции подстановки и перестановки, обеспечивающие рассеивание и перемешивание информации в сообщениях. Сообщения в подавляющем своем большинстве обладают информационной избыточностью и подвержены статистическому анализу, который в совокупности с методом вероятных слов существенно облегчает жизнь злоумышленникам.

В свою очередь, при информационном подходе отмечается, что информация в широком смысле – это фундаментальное свойство материи, при котором происходит отображение одного множества на другое, при этом состояние одного множества находится в соответствии с состоянием другого множества. Существует два основных направления использования информационного подхода в теории систем и системном анализе [3-5], это:

1) отображение информации о системах, количественное измерение информационных потоков, имеющее большое значение для моделирования систем [7,11,15];

2) процессы управления и преобразования, происходящие в системах, для их понимания используются такие основные понятия теории информации как неопределенность, энтропия и количество информации [7,11,16].

В ряде работ при информационном подходе рассматривается теория информационного поля, разработанная Денисовым А.А., на ее основе были промоделированы процессы в физически неоднородных системах, некоторые экономические явления и др. [5,19-22].

По происхождению системы, как правило, разделяются на естественные, искусственные и смешанные системы.

Искусственные системы, к которым относятся и криптографические, подразделяются на:

1) материальные, состоящие из физических объектов, собранных человеком в систему с определенной целью (конкретные криптографические алгоритмы, системы и функции, например, DES, AES, Люцифер, ГОСТ 28.147-89 и др.), и

2) абстрактные системы, созданные средствами мышления. Абстрактные системы по существу являются моделями, которые могут быть аналитическими, логическими, алгоритмическими, образными и т.д. Абстрактные системы, как правило, имеют и материальное воплощение, которое выражается в виде текста, формул, диаграмм, рисунков и т.д.

Системы могут также классифицироваться также по сложности, степени изолированности, по характеру функционирования, степени организованности, по способу задания целей и управления [3-5].

В данной работе рассмотрены виды систем, а также приведены различные варианты определения систем, в частности, криптографических, в рамках понятийного аппарата общей теории систем и системного анализа. При этом авторами проанализирован и дополнен набор свойств систем по сравнению с аналогичной классификацией, приведенной в [6]. Дополнительно уточнены наименования подходов в рамках теории систем. Также авторами проанализирован и представлен дополненный и уточненный по сравнению с [3-5] жизненный цикл данных систем.

Библиографический список

1. Волкова В.Н., Денисов А.А. Основы теории систем и системного анализа: учебник для студентов вузов. – Изд. 2-е перераб. и доп. – СПб.: Изд-во СПбГТУ, 1999. – 512 с.
2. Садовский В.Н. Системный анализ // Новая философская энциклопедия: В 4 т. Т. III.
3. Павлов С.Н. Теория систем и системный анализ: Учебное пособие. - Томск: Томский межвузовский центр дистанционного образования, 2003. – 134 с.
4. Вдовин В.М. теория систем и системный анализ: Учебник/ В.М. Вдовин, Л.Е. суркова, В.А. Валентинов. - М.: издательско-торговая корпорация «Дашков и к», 2010. – 134 с.
5. Силич Мария Петровна. Основы теории систем и системного анализа: учебное пособие/М.П. Силич, В.А. Силич. - Томск: Гос ун-та систем управления и радиоэлектроники, 2013. – 340 с.
6. Алферов А.П., Зубов В.Ю., Кузьмин А.С., Черемушкин А.В. Основы криптографии: Учебное пособие. - М.: Гелиос АРН, 2001. – 480 с., ил.
7. Мартынов А.П., Фомченко В.Н. Криптография и электроника/ Под ред. А.И. Астайкина. Саров: ФГУП «РФЯЦ-ВНИИЭФ», 2006. - 452с.
8. Кузин Л.Т. Т.1. Математические основы кибернетики: Учебное пособие для студентов вузов. - М.: «Энергия», 1973. – 505 с., ил.

9. Бергаланфи Л. фон. Общая теория систем – критический обзор // Исследования по общей теории систем: Сб. переводов. М.: Прогресс, 1969.
10. Мартынов А.П., Мартынова И.А., Фомченко В.Н. Аксиоматические основы функций подстановки в системе счисления ряда факториальных множеств и их характеристики: Монография. – Саров: ФГУП «РФЯЦ-ВНИИЭФ», 2019. – 210с., ил.
11. Мартынова И.А., Машин И.Г., Фомченко В.Н. Теория поля и защита информации: Монография. – Саров: ФГУП «РФЯЦ-ВНИИЭФ», 2017. – 209с., ил.
12. Уемов А.И. Системный подход и общая теория систем. - М.: Мысль, 1978. – 272 с.
13. Грибунин В.Г., Костюков В.Е., Мартынов А.П., Николаев Д.Б., Фомченко В.Н. Современные методы обеспечения безопасности в атомной энергетике: Монография/ под ред. А.И. Астайкина. Саров: ФГУП «РФЯЦ-ВНИИЭФ», 2014. – 636с., ил.
14. Защита информации. Исследование общетеоретической концепции и архитектурных принципов построения распределенных адаптивных систем с функцией защиты данных: Учебное пособие / И.А. Мартынова, С.В. Смуров, П.В. Воронов, И.А. Николаева. - Серпухов: ФВА РВСН имени Петра Великого, 2017. – 176 с.
15. Обеспечение подлинности и целостности информации при проведении научных экспериментов: Учебное пособие / И.А. Мартынова, С.В. Смуров, П.В. Воронов, И.А. Николаева. - Серпухов: ФВА РВСН имени Петра Великого, 2016. – 226 с.
16. Мартынова И.А., Машин И.Г., Фомченко В.Н. Введение в теорию поля и ее приложения: Монография. – Саров: ФГУП «РФЯЦ-ВНИИЭФ», 2014. – 108 с., ил.
17. Кан Д. Взломщики кодов: Пер. с англ. А Ключевского. - М.: ЗАО Центр-Полиграф, 2000.
18. Риксон, Фред Б. Коды, шифры, сигналы и тайная передача информации: Пер. с англ. А. Галыгина. - М.: АСТ: Астрель, 2003.
19. Пат. 2256053. Устройство управления электронными замками / Гончаров С.Н., Мартынов А.П., Николаев Д.Б., Фомченко В.Н. Заявл. 20.01.2003, опубл. 10.07.2005. Бюл. № 19.
20. Пат. 2224075. Электронно-механический замок / Мартынов А.П., Николаев Д.Б., Сычев В.Т., Шишкин Г.И. Заявл. 28.11.2001, опубл. 20.02.2004. Бюл. № 5.
21. Пат. 2154722. Электронный кодовый замок / Мартынов А.П., Николаев Д.Б., Шишкин Г.И. Заявл. 27.08.1999, опубл. 20.08.2000. Бюл. № 23.
22. Грибунин В.Г., Костюков В.Е., Мартынов А.П., Николаев Д.Б., Фомченко В.Н. Стеганографические системы. Атаки, пропускная способность каналов и оценка стойкости: Учебно-методическое пособие // Под ред. В.Г. Грибунина. Саров: ФГУП "РФЯЦ-ВНИИЭФ", 2015.

I. A. Martynova¹, N.V. Gorbatenko², O.V. Krivosheev²

THE TYPES AND THE DEFINITIONS OF THE CRYPTOGRAPHIC SYSTEMS

¹ *Sarov Institute of Physics and Technology - branch of NRNU MEPhI*

² *Russian Federal Nuclear Center –*

All-Russian Scientific Research Institute of Experimental Physics (RFNC-VNIIEF)

The article considers the types of systems and various options for defining systems, in particular, cryptographic systems, within the framework of the conceptual apparatus of the general theory of systems. It also offers an expanded and refined life cycle of these systems.

Keywords: system, system approach, system theory, types of systems, life cycle of systems.

**ИНТЕРПРЕТАЦИОННЫЙ ПОДХОД ДЛЯ МЕТОДИКИ АНАЛИЗА
ХАРАКТЕРИСТИК МАТЕМАТИЧЕСКИХ ОПЕРАЦИЙ**

Саровский физико-технический институт – филиал НИЯУ МИФИ

Дано описание интерпретационного подхода для формирования методики проведения аналитических исследований математических операций на примере анализа элементов ряда факториальных множеств, проведено исследование факторов разложения рядов факториальных множеств при применении современных математических преобразований информации.

Ключевые слова: ряд факториального множества, декремент, инверсия, обратная функция, умножение множеств, коммутанты.

В основе любых криптографических алгоритмов и протоколов используются простейшие функции – перестановки и подстановки элементов заданного конечного множества. Объединяя и трансформируя данные элементы, организуется структурная ячейка защиты информации – математическое преобразование, обладающее необходимой криптографической стойкостью. Главная проблема в выработке таких преобразованиях заключается в том, что в процессе формирования более мощных алгоритмов разработчики используют все больше данных и вычислительных мощностей, полагаясь при этом на аппаратные возможности информационной системы. Это не только увеличивает потребление ресурсов, но также ограничивает скорость и конфиденциальность приложений.

Система счисления ряда факториальных множеств позволяет использовать интерпретационный подход проведения аналитических исследований математических операций посредством формирования любого элемента факториального множества без хранения перестановок в оперативной памяти. Таким образом, функция хранения массивов данных для криптографических алгоритмов уже не является необходимостью, так как ее функциональные возможности заменяет использование алгоритма формирования перестановок из системы счисления рядов факториальных множеств [1].

В работе [2] предложена система счисления ряда факториальных множеств, представлены способы преобразования чисел из десятичной системы счисления в систему счисления факториальных множеств и обратно, обеспечивающие обратимое и взаимно однозначное преобразование и нумерацию элементов факториальных множеств любой размерности. Предложен способ преобразования образов ряда факториальных множеств в конкретные перестановки, имеющий большое значение для теории защиты информации и криптографии.

Для оптимизации работы с системой счисления факториальных множеств разработан «Программный комплекс анализа подстановок ряда факториальных множеств» [3], который формирует исходные данные для дальнейшего анализа критериальных факторов разложения факториальных множеств. Основные возможности комплекса:

- обеспечение подсчета перестановки по ее номеру из факториальной системы счисления;
- обеспечение расчета всего множества и его критериальных факторов (циклическая форма, декремент, инверсия, обратная функция);
- табличное умножение конечных факториальных множеств (также и коммутантов) с выбором элементов (строка и столбец) для перемножения;

- табличное умножение конечных факториальных множеств (также и коммутантов) с выборочным выводом – перестановка, циклическая форма, декремент или по номеру.

Схема-алгоритм работы программного комплекса представлен на рис. 1.

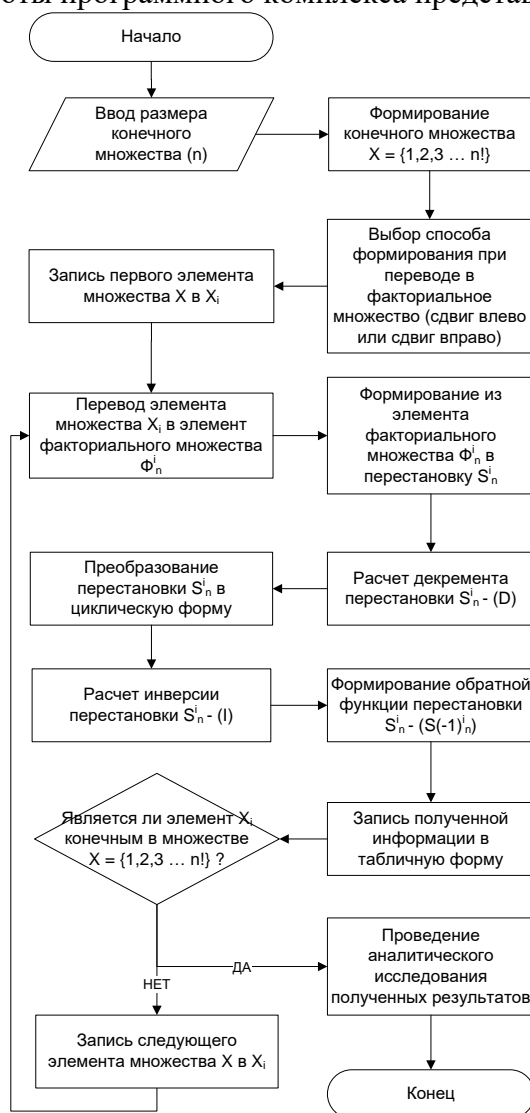


Рис. 1. Схема-алгоритм работы программного комплекса подстановок ряда факториальных множеств

Немного вернемся к способу формирования перестановок из факториальной системы. В работе [2] предложен способ трансформирования из множества Φ_n факториальной системы счисления в перестановку сдвигом вправо каждого последующего элемента в регистре сдвига. Имея такой механизм – из любого элемента множества Φ_n формируется новый элемент – S перестановка, притом уникальная для данного конечного множества Φ_n .

Попробуем проделать такой же механизм, но только со сдвигом влево элемента в регистре сдвига. Сравнение двух механизмов формирования перестановки представлены на рис. 2 и 3.

Стоит отметить, что результаты разные, и можно считать, что, используя при формировании перестановок разные способы сдвига элементов – на выходе получаются разные множества S_n , которые обладают разными свойствами.

Пример выходных данных программного комплекса для анализа подстановок ряда факториальных множеств показан на рис. 4.

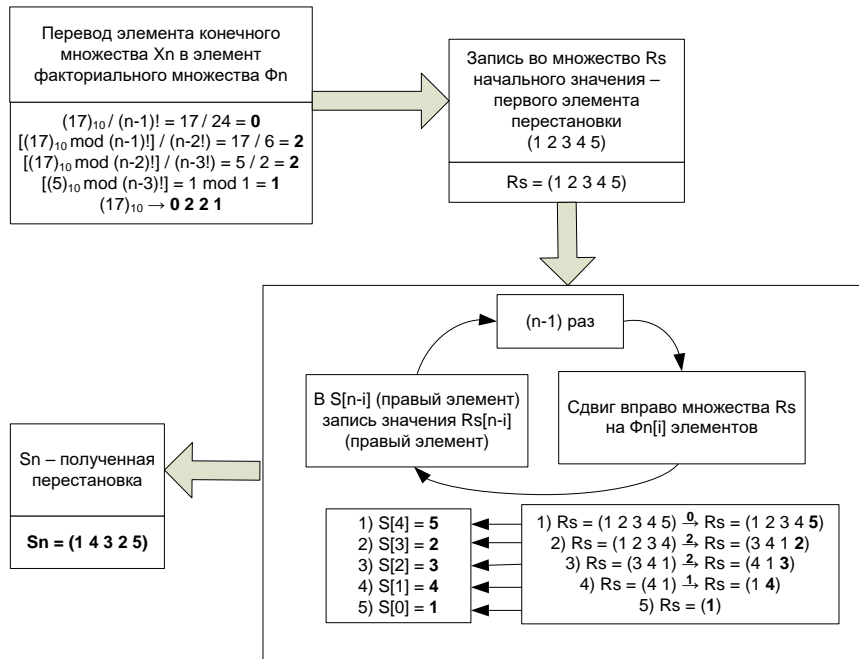


Рис. 2. Формирование перестановки сдвигом вправо

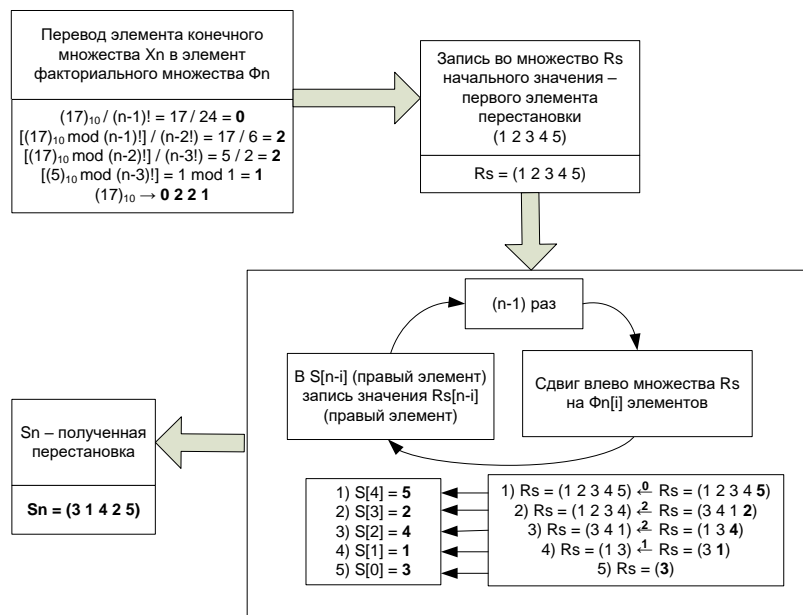


Рис. 3. Формирование перестановки сдвигом влево

Следует обратить внимание на третий и девятый столбец – это уже сформированные множества перестановок S_i , и они разные. Следовательно, что тип формирования подстановки из факториальной системы счисления (то есть сдвигом вправо или сдвигом влево) позволяет формировать разные множества.

Далее проведено исследование факторов разложения рядов факториальных множеств при применении современных математических преобразований информации, выработано техническое решение о внедрении рядов факториальных множеств в процесс защиты информационных систем.

	1	2	3	4	5	6	7	8	9	10	11	12	13	14
1	X	Ф	S	Цикл. форма	D	I	S(-1)	№ S(-1)	S лево	Цикл. форма	D	I	S(-1)	№ S(-1)
2	0	0000	12345	(1)(2)(3)(4)(5)	0	0	12345	0	12345	(1)(2)(3)(4)(5)	0	0	12345	0
3	1	0001	21345	(12)(3)(4)(5)	1	1	21345	1	21345	(12)(3)(4)(5)	1	1	21345	1
4	2	0010	31245	(132)(4)(5)	2	2	23145	4	23145	(123)(4)(5)	2	2	31245	4
5	3	0011	13245	(1)(23)(4)(5)	1	1	13245	3	32145	(13)(2)(4)(5)	1	3	32145	3
6	4	0020	23145	(123)(4)(5)	2	2	31245	2	31245	(132)(4)(5)	2	2	23145	2
7	5	0021	32145	(13)(2)(4)(5)	1	3	32145	5	13245	(1)(23)(4)(5)	1	1	13245	5
8	6	0100	41235	(1432)(5)	3	3	23415	18	23415	(1234)(5)	3	3	41235	18
9	7	0101	14235	(1)(243)(5)	2	2	13425	14	32415	(134)(2)(5)	2	4	42135	23
10	8	0110	24135	(1243)(5)	3	3	31425	15	34215	(1324)(5)	3	5	43125	13
11	9	0111	42135	(143)(2)(5)	2	4	32415	19	43215	(14)(23)(5)	2	6	43215	9
12	10	0120	12435	(1)(2)(34)(5)	1	1	12435	10	42315	(14)(2)(3)(5)	1	5	42315	10
13	11	0121	21435	(12)(34)(5)	2	2	21435	11	24315	(124)(3)(5)	2	4	41325	14
14	12	0200	34125	(13)(24)(5)	2	4	34125	12	34125	(13)(24)(5)	2	4	34125	12
15	13	0201	43125	(1423)(5)	3	5	34215	22	43125	(1423)(5)	3	5	34215	8
16	14	0210	13425	(1)(234)(5)	2	2	14235	7	41325	(142)(3)(5)	2	4	24315	11
17	15	0211	31425	(1342)(5)	3	3	24135	8	14325	(1)(24)(3)(5)	1	3	14325	15
18	16	0220	41325	(142)(3)(5)	2	4	24315	21	13425	(1)(234)(5)	2	2	14235	19
19	17	0221	14325	(1)(24)(3)(5)	1	3	14325	17	31425	(1342)(5)	3	3	24135	22
20	18	0300	23415	(1234)(5)	3	3	41235	6	41235	(1432)(5)	3	3	23415	6
21	19	0301	32415	(134)(2)(5)	2	4	42135	9	14235	(1)(243)(5)	2	2	13425	16
22	20	0310	42315	(14)(2)(3)(5)	1	5	42315	20	12435	(1)(2)(34)(5)	1	1	12435	20
23	21	0311	24315	(124)(3)(5)	2	4	41325	16	21435	(12)(34)(5)	2	2	21435	21

Рис. 4. Пример выходных данных программного комплекса для анализа подстановок ряда факториальных множеств

Рассмотрены критериальные факторы разложения рядов факториальных множеств, такие как декремент, инверсия, обратная функция, в дальнейшем исследовании будет проводиться доказательство данного способа в использовании современных стандартов в области защиты информации и высокотехнологичных технических средств защиты информации.

Рассмотрим факторы разложения рядов факториальных множеств:

1) Декремент – фактор, описывающий разность числа всех индексов ряда факториального множества и количества циклов, включая циклы единичной длины (пятый и одиннадцатый столбцы рис. 4);

2) Инверсия – фактор, определяющий суммарное количество элементов ряда факториального множества, которые располагаются правее его меньших индексов (шестой и двенадцатый столбцы рис. 4);

3) Обратная функция ряда факториального множества – фактор, определяющий новый ряд факториального множества за счет построения обратной функции множества (столбцы № 7, 8, 13, 14 рис. 4).

Опираясь на исследуемые факторы, в дальнейшем можно описывать концептуальные особенности задания рядов факториального множества и применять их в теоретических основах построения систем защиты информации. Предложенные факторы разложения рядов факториального множества позволяют использовать построенный на их основе математический аппарат в процессе преобразования информации при реализации современных стандартов и построении высокотехнологичных технических средств защиты информации, так как с помощью рядов происходит «перемешивание» элементов множества с заданными характеристиками.

Функциональные возможности «Программного комплекса анализа подстановок ряда факториальных множеств» [3,4] позволяют совершать умножение конечных множеств для дальнейшего использования уже получившихся на выходе операции множеств.

Пример умножения множества подстановок при формировании сдвигом вправо представлен на рис. 5, а умножения множества коммутантов подстановок при формировании сдвигом влево (вывод циклической формы) на рис. 6.

В ходе проведенных работ реализован программный комплекс анализа подстановок ряда факториальных множеств, который использует предложенный авторами интерпретационный подход и реализован на кроссплатформенном фреймворке Q для разработки программного обеспечения на языке программирования C++. Данный комплекс позволяет решать задачи по оптимизации работы со системой счисления ряда факториальных множеств, а также по формированию множеств перестановок, которые будут использоваться в адаптивных алгоритмах преобразования информации [5-11].

1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	
1		1234	2134	3124	1324	2314	3214	4123	1423	2413	4213	1243	2143	3412	4312	1342	3142	4132	1432	2341	3241	4231	2431	3421	4321
2	1234	1234	2134	3124	1324	2314	3214	4123	1423	2413	4213	1243	2143	3412	4312	1342	3142	4132	1432	2341	3241	4231	2431	3421	4321
3	2134	2134	1234	1324	3124	3214	2314	1423	4123	4213	2413	2143	1243	1243	4312	3412	3142	1342	1432	4132	3241	2341	2431	4231	3421
4	3124	3124	3214	2314	2134	1234	1324	2413	2143	1243	1423	4123	4213	3412	4312	4132	3142	4132	1342	4231	3241	4321	3421	2341	2431
5	1324	1324	2413	3412	1432	2341	3241	3124	4213	1243	4123	1423	2143	3142	4132	1432	4312	4132	1342	2431	3421	4321	2341	3421	4231
6	2314	2314	1324	1234	3214	3124	2134	1243	4213	4123	2143	2413	1423	4132	3142	3412	1432	1342	4312	3421	2431	2341	4321	4231	3241
7	3214	3214	3124	2134	2314	1324	1234	2143	2413	1423	1243	4213	4123	1432	1342	4312	4132	3142	4132	3142	4321	4231	3241	3421	2341
8	4123	4123	4213	4312	4132	4231	4321	3412	3142	3241	3421	3124	3214	2341	2431	2134	2314	2413	2143	1234	1324	1423	1243	1342	1432
9	1423	1423	2413	3412	1432	2431	3421	4312	1342	2341	4321	1324	2314	3241	4231	1234	3214	4213	1243	2134	3124	4123	2143	3142	4132
10	2413	2413	1423	1432	3412	3421	2431	1342	4312	4321	2341	2314	1324	4231	3241	3214	1234	1243	4213	3124	2134	2143	4123	4132	3142
11	4213	4213	4123	4132	4312	4321	4231	3142	3412	3421	3241	3214	3124	2431	2341	2314	2134	2143	2413	1324	1234	1243	1423	1432	1342
12	1243	1243	2143	3142	1342	2341	3241	4132	1432	2431	4231	1234	2134	3421	4321	1324	3124	4123	1423	2314	3214	4213	2143	3412	4312
13	2143	2143	1243	1342	3142	3241	2341	1432	4132	4231	2431	2134	1234	4321	3421	3124	1324	1423	4123	3214	2314	2413	4213	4132	3412
14	3412	3412	3421	2431	2413	1423	1432	2341	2314	1324	1342	4312	4321	1234	1243	4213	4231	3241	3214	4123	4132	3142	3124	2134	2143
15	4312	4312	4321	4231	4213	4123	4132	3241	3214	3124	3142	3412	3421	2134	2143	2413	2431	2341	2314	1423	1432	1342	1324	1234	1243
16	1342	1342	2341	3241	1243	2143	3142	4231	1234	2134	4132	1432	2431	3124	4123	1423	3421	4321	1324	2413	3412	4312	2314	3214	4213
17	3142	3142	3241	2341	2143	1243	1342	2431	2134	1234	1432	4132	4231	1324	1423	4123	4321	3421	3124	4213	4312	3412	3214	2314	2413
18	4132	4132	4231	4321	4123	4132	4213	3412	3124	3214	3412	3421	2314	2413	2143	2341	2431	2134	2143	1342	1432	1324	1234	1324	1423
19	1432	1432	2431	3421	1423	2413	3412	4321	1324	2314	4312	1342	2341	3214	4213	1243	3241	4231	1234	2143	3142	4132	2134	3124	4123
20	2341	2341	1342	1243	3241	3142	2143	1234	4231	4132	2134	2431	1432	4123	3124	3421	1423	1324	4321	3412	2413	2314	4312	4213	3214
21	3241	3241	3142	2143	2341	1342	1243	2134	2431	1432	1234	4231	4132	1423	1423	1324	4321	4123	3124	3421	4312	4213	3214	3412	2413
22	4231	4231	4132	4123	4321	4312	4213	3124	3421	3412	3214	3241	3142	2413	2314	2341	2143	2134	2431	1342	1243	1234	1432	1423	1324
23	2431	2431	1432	1423	3421	3412	2413	1324	4321	4312	2314	2341	1342	4213	3214	3241	1243	1234	4231	3142	2143	2134	4132	4123	3124
24	3421	3421	3412	2413	2431	1432	1423	2314	2341	1342	1324	4321	4312	1243	1234	4231	4213	3214	3241	4132	4123	3124	3142	2143	2134
25	4321	4321	4312	4213	4231	4132	4123	3214	3241	3142	3124	3421	3412	2143	2134	2431	2413	2314	2341	1432	1423	1324	1342	1243	1234

Рис. 5. Умножение множества подстановок при формировании сдвигом вправо

1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19
1	1234	2134	2314	3214	3124	1324	2341	3241	3421	4321	4231	2431	3412	4312	4132	1432	1342	3142
2	1234	(1)(2)(3)(4)	(1)(2)(3)(4)	(1)(2)(3)(4)	(1)(2)(3)(4)	(1)(2)(3)(4)	(1)(2)(3)(4)	(1)(2)(3)(4)	(1)(2)(3)(4)	(1)(2)(3)(4)	(1)(2)(3)(4)	(1)(2)(3)(4)	(1)(2)(3)(4)	(1)(2)(3)(4)	(1)(2)(3)(4)	(1)(2)(3)(4)	(1)(2)(3)(4)	(1)(2)(3)(4)
3	2134	(1)(2)(3)(4)	(1)(2)(3)(4)	(1)2(3)(4)	(1)3(2)(4)	(1)3(2)(4)	(1)2(3)(4)	(1)2(4)(3)	(1)4(2)(3)	(1)2(3)(4)	(1)2(3)(4)	(1)4(2)(3)	(1)2(4)(3)	(1)2(4)(3)	(1)2(4)(3)	(1)2(4)(3)	(1)2(4)(3)	(1)2(4)(3)
4	2314	(1)(2)(3)(4)	(1)3(2)(4)	(1)(2)(3)(4)	(1)3(2)(4)	(1)(2)(3)(4)	(1)3(2)(4)	(1)(2)3(4)	(1)4(2)3	(1)3(2)(4)	(1)3(2)(4)	(1)3(2)(4)	(1)3(2)(4)	(1)3(2)(4)	(1)3(2)(4)	(1)3(2)(4)	(1)3(2)(4)	(1)3(2)(4)
5	3214	(1)(2)(3)(4)	(1)2(3)(4)	(1)3(2)(4)	(1)3(2)(4)	(1)(2)(3)(4)	(1)3(2)(4)	(1)3(2)(4)	(1)3(2)(4)	(1)3(2)(4)	(1)3(2)(4)	(1)3(2)(4)	(1)3(2)(4)	(1)3(2)(4)	(1)3(2)(4)	(1)3(2)(4)	(1)3(2)(4)	(1)3(2)(4)
6	3124	(1)(2)(3)(4)	(1)2(3)(4)	(1)(2)(3)(4)	(1)2(3)(4)	(1)(2)(3)(4)	(1)2(3)(4)	(1)3(4)(2)	(1)3(2)(4)	(1)2(4)3	(1)2(3)(4)	(1)4(2)3	(1)2(3)(4)	(1)4(2)3	(1)4(2)3	(1)4(2)3	(1)4(2)3	(1)4(2)3
7	1324	(1)(2)(3)(4)	(1)3(2)(4)	(1)2(3)(4)	(1)2(3)(4)	(1)3(2)(4)	(1)3(2)(4)	(1)2(3)(4)	(1)2(3)(4)	(1)2(3)(4)	(1)2(3)(4)	(1)2(3)(4)	(1)2(3)(4)	(1)2(3)(4)	(1)2(3)(4)	(1)2(3)(4)	(1)2(3)(4)	(1)2(3)(4)
8	2341	(1)(2)(3)(4)	(1)4(2)(3)	(1)(2)4(3)	(1)3(2)(4)	(1)4(3)(2)	(1)3(2)(4)	(1)(2)(3)(4)	(1)4(2)(3)	(1)(2)4(3)	(1)3(2)(4)	(1)4(3)(2)	(1)3(2)(4)	(1)4(3)(2)	(1)4(3)(2)	(1)4(3)(2)	(1)4(3)(2)	(1)4(3)(2)
9	3241	(1)(2)(3)(4)	(1)2(4)(3)	(1)4(2)(3)	(1)4(3)(2)	(1)3(2)(4)	(1)3(2)(4)	(1)3(2)(4)	(1)(2)(3)(4)	(1)3(2)(4)	(1)3(2)(4)	(1)4(3)(2)	(1)3(2)(4)	(1)4(3)(2)	(1)4(3)(2)	(1)4(3)(2)	(1)4(3)(2)	(1)4(3)(2)
10	3421	(1)(2)(3)(4)	(1)2(3)(4)	(1)4(2)(3)	(1)4(3)(2)	(1)2(3)(4)	(1)2(3)(4)	(1)2(3)(4)	(1)(2)(3)(4)	(1)2(3)(4)	(1)2(3)(4)	(1)4(2)(3)	(1)4(3)(2)	(1)2(3)(4)	(1)2(3)(4)	(1)2(3)(4)	(1)2(3)(4)	(1)2(3)(4)
11	4321	(1)(2)(3)(4)	(1)2(3)(4)	(1)3(2)(4)	(1)3(2)(4)	(1)2(3)(4)	(1)3(2)(4)	(1)3(2)(4)	(1)3(2)(4)	(1)3(2)(4)	(1)3(2)(4)	(1)3(2)(4)	(1)3(2)(4)	(1)3(2)(4)	(1)3(2)(4)	(1)3(2)(4)	(1)3(2)(4)	(1)3(2)(4)
12	4231	(1)(2)(3)(4)	(1)2(4)(3)	(1)3(4)(2)	(1)3(4)(2)	(1)2(4)(3)	(1)(2)(3)(4)	(1)3(4)(2)	(1)3(4)(2)	(1)2(4)(3)	(1)(2)(3)(4)	(1)(2)(3)(4)	(1)2(4)(3)	(1)2(4)(3)	(1)2(4)(3)	(1)2(4)(3)	(1)2(4)(3)	(1)2(4)(3)
13	2431	(1)(2)(3)(4)	(1)4(2)(3)	(1)2(3)(4)	(1)3(4)(2)	(1)4(2)(3)	(1)2(3)(4)	(1)4(2)(3)	(1)4(2)(3)	(1)4(2)(3)	(1)4(2)(3)	(1)4(2)(3)	(1)4(2)(3)	(1)4(2)(3)	(1)4(2)(3)	(1)4(2)(3)	(1)4(2)(3)	(1)4(2)(3)
14	3412	(1)(2)(3)(4)	(1)2(3)(4)	(1)2(3)(4)	(1)(2)(3)(4)	(1)4(2)(3)	(1)4(2)(3)	(1)4(2)(3)	(1)2(3)(4)	(1)2(3)(4)	(1)2(3)(4)	(1)4(2)(3)	(1)4(2)(3)	(1)4(2)(3)	(1)4(2)(3)	(1)4(2)(3)	(1)4(2)(3)	(1)4(2)(3)
15	4312	(1)(2)(3)(4)	(1)2(3)(4)	(1)3(4)(2)	(1)3(2)(4)	(1)2(4)(3)	(1)(2)4(3)	(1)2(4)(3)	(1)2(4)(3)	(1)2(4)(3)	(1)2(4)(3)	(1)2(4)(3)	(1)2(4)(3)	(1)2(4)(3)	(1)2(4)(3)	(1)2(4)(3)	(1)2(4)(3)	(1)2(4)(3)
16	4132	(1)(2)(3)(4)	(1)2(4)(3)	(1)3(2)(4)	(1)3(2)(4)	(1)2(3)(4)	(1)(2)3(4)	(1)2(3)(4)	(1)2(3)(4)	(1)2(3)(4)	(1)2(3)(4)	(1)2(3)(4)	(1)2(3)(4)	(1)2(3)(4)	(1)2(3)(4)	(1)2(3)(4)	(1)2(3)(4)	(1)2(3)(4)
17	1432	(1)(2)(3)(4)	(1)4(2)(3)	(1)4(2)(3)	(1)(2)(3)(4)	(1)2(3)(4)	(1)2(3)(4)	(1)3(2)(4)	(1)(2)4(3)	(1)(2)4(3)	(1)3(2)(4)	(1)2(4)(3)	(1)2(4)(3)	(1)2(4)(3)	(1)2(4)(3)	(1)2(4)(3)	(1)2(4)(3)	(1)2(4)(3)
18	1342	(1)(2)(3)(4)	(1)4(2)(3)	(1)4(2)(3)	(1)2(3)(4)	(1)3(2)(4)	(1)(2)4(3)	(1)3(4)(2)	(1)2(3)(4)	(1)3(4)(2)	(1)2(3)(4)	(1)3(4)(2)	(1)2(3)(4)	(1)3(4)(2)	(1)3(4)(2)	(1)3(4)(2)	(1)3(4)(2)	(1)3(4)(2)
19	3142	(1)(2)(3)(4)	(1)2(4)(3)	(1)2(4)(3)	(1)4(3)(2)	(1)2(3)(4)	(1)4(3)(2)	(1)2(3)(4)	(1)2(4)(3)	(1)2(4)(3)	(1)2(4)(3)	(1)2(4)(3)	(1)2(4)(3)	(1)2(4)(3)	(1)2(4)(3)	(1)2(4)(3)	(1)2(4)(3)	(1)2(4)(3)
20	4123	(1)(2)(3)(4)	(1)2(3)(4)	(1)3(4)(2)	(1)3(2)(4)	(1)2(4)(3)	(1)(2)3(4)	(1)2(3)(4)	(1)2(3)(4)	(1)2(3)(4)	(1)2(3)(4)	(1)2(3)(4)	(1)2(3)(4)	(1)2(3)(4)	(1)2(3)(4)	(1)2(3)(4)	(1)2(3)(4)	(1)2(3)(4)
21	1423	(1)(2)(3)(4)	(1)3(2)(4)	(1)2(3)(4)	(1)4(3)(2)	(1)4(2)(3)	(1)(2)3(4)	(1)2(4)(3)	(1)3(2)(4)	(1)2(4)(3)	(1)3(2)(4)	(1)2(4)(3)	(1)2(4)(3)	(1)2(4)(3)	(1)2(4)(3)	(1)2(4)(3)	(1)2(4)(3)	(1)2(4)(3)
22	1243	(1)(2)(3)(4)	(1)(2)(3)(4)	(1)(2)4(3)	(1)4(3)(2)	(1)4(3)(2)	(1)(2)4(3)	(1)2(3)(4)	(1)3(4)(2)	(1)2(3)(4)	(1)3(4)(2)	(1)2(3)(4)	(1)3(4)(2)	(1)2(3)(4)	(1)2(3)(4)	(1)2(3)(4)	(1)2(3)(4)	(1)2(3)(4)
23	2143	(1)(2)(3)(4)	(1)(2)(3)(4)	(1)4(2)(3)	(1)3(2)(4)	(1)3(2)(4)	(1)3(2)(4)	(1)3(2)(4)	(1)3(2)(4)	(1)3(2)(4)	(1)3(2)(4)	(1)3(2)(4)	(1)3(2)(4)	(1)3(2)(4)	(1)3(2)(4)	(1)3(2)(4)	(1)3(2)(4)	(1)3(2)(4)
24	2413	(1)(2)(3)(4)	(1)3(2)(4)	(1)4(2)(3)	(1)3(4)(2)	(1)(2)3(4)	(1)4(2)(3)	(1)3(4)(2)	(1)4(2)(3)	(1)3(2)(4)	(1)(2)(3)(4)	(1)4(2)(3)	(1)4(2)(3)	(1)4(2)(3)	(1)4(2)(3)	(1)4(2)(3)	(1)4(2)(3)	(1)4(2)(3)
25	4213	(1)(2)(3)(4)	(1)2(3)(4)	(1)3(2)(4)	(1)3(4)(2)	(1)2(3)(4)	(1)(2)4(3)	(1)2(3)(4)	(1)(2)(3)(4)	(1)(2)4(3)	(1)2(3)(4)	(1)2(4)(3)	(1)2(4)(3)	(1)2(4)(3)	(1)2(4)(3)	(1)2(4)(3)	(1)2(4)(3)	(1)2(4)(3)

Рис. 6. Умножение множества коммутантов подстановок при формировании сдвигом влево (вывод циклической формы)

Библиографический список

1. Мартынов А.П., Мартынова И.А., Фомченко В.Н. Аксиоматические основы функций подстановки в системе счисления ряда факториальных множеств и их характеристики. Монография. – Саров: ФГУП «РФЯЦ-ВНИИЭФ», 2019. 210 с.
2. Мартынов А.П., Мартынова И.А. Функции перестановки в системе счисления ряда факториальных множеств. Вестник ВГУ, Серия: Системный анализ и информационные технологии, № 3, 2016, стр. 42-49
3. Свидетельство о государственной регистрации программы для ЭВМ № 2020613795. Программный комплекс анализа подстановок ряда факториальных множеств / Мартынов А.П., Николаев Д.Б., Сплюхин Д.В., Фомченко В.Н., Мартынова И.А. Зарег. 23.03.2020 г. – М.: Роспатент, 2020
4. Мартынов А.П., Николаев Д.Б., Мартынова И.А. Криптографические системы и метод факториального сжатия информации. – Известия института инженерной физики, 2016. № 4(42). С.54-57.
5. Пат. 2408924, 10.01.2011. Устройство обеспечения безопасного формирования параметров / Курочкин А.А., Мартынов А.П., Николаев Д.Б., Фомченко В.Н. Патент на изобретение № 2408924 от 10.01.2011.
6. Пат. 2623894, 07.09.2017. Способ преобразования данных с равновероятностной инициализацией / Мартынов А.П., Мартынова И.А., Марунин М.В., Николаев Д.Б., Фомченко В.Н. Патент на изобретение № 2623894 от 07.09.2017.
7. Пат. 2546238, 07.10.2013. Способ транслитерационного преобразования информации и передачи ее по каналам связи / Грибунин В.Г., Мартынов А.П., Николаев Д.Б., Фомченко В.Н. Патент на изобретение № 2546238 от 07.10.2013.
8. Пат. 2554525, 07.10. 2013.Способ преобразования информации с синхронной сменой инициализирующих последовательностей в блоках, соединенных каналом связи с неопределенным периодом смены / Фомченко В.Н., Мартынов А.П., Николаев Д.Б., Масягин А.М., Мартынов А.А. Патент на изобретение № 2554525 от 07.10.2013.
9. Пат. 2256053. Устройство управления электронными замками / Гончаров С.Н., Мартынов А.П., Николаев Д.Б., Фомченко В.Н. Заявл. 20.01.2003, опубл. 10.07.2005. Бюл. № 19.
10. Пат. 2224075. Электронно-механический замок / Мартынов А.П., Николаев Д.Б., Сычев В.Т., Шишкин Г.И. Заявл. 28.11.2001, опубл. 20.02.2004. Бюл. № 5.
11. Пат. 2154722. Электронный кодовый замок / Мартынов А.П., Николаев Д.Б., Шишкин Г.И. Заявл. 27.08.1999, опубл. 20.08.2000. Бюл. № 23.

I.A. Martynova

INTERPRETATION APPROACH FOR ANALYSIS OF THE CHARACTERISTICS OF MATHEMATICAL OPERATIONS

Sarov Institute of Physics and Technology - branch of NRNU MEPhI

The description of an interpretive approach for the formation of a methodology for conducting analytical research of mathematical operations on the example of the analysis of elements of a number of factorial sets is given, a study of the factors of expansion of series of factorial sets when applying modern mathematical transformations of information is carried out.

Keywords: series of factorial set, decrement, inversion, inverse function, multiplication of sets, commutants.

**МЕТОДОЛОГИЯ НАУЧНОГО ПОЗНАНИЯ НА ПРИМЕРЕ ФОРМИРОВАНИЯ
АКСИОМАТИЧЕСКОГО ПОДХОДА В ЭЛЕМЕНТАРНОЙ АЛГЕБРЕ***Саровский физико-технический институт – филиал НИЯУ МИФИ*

Дано описание интерпретационного подхода для аналитических исследований элементов ряда факториальных множеств, проведено исследование факторов разложения рядов факториальных множеств при применении современных математических преобразований информации.

Ключевые слова: ряд факториального множества, декремент, инверсия, обратная функция, умножение множеств, коммутанты.

Стремительное увеличение объемов информации и переход к цифровой экономике как средству достижения поставленной цели, привели к бурному развитию информационных технологий и необходимости значительной перестройке информационной среды. Это не только открывает новые возможности научно-технического и технологического прогресса, но и добавляет новые угрозы, связанные с глобализацией информации и необходимостью ее защиты, которая в основном обеспечивается криптографическими методами. Все это повышает актуальность более глубокого анализа и изучения существующих криптографических функций, алгоритмов, протоколов и систем, и делает актуальным поиск и создание новых методов, функций и алгоритмов защиты информации.

Весь исторический опыт развития криптографии [1-3] показывает, что основными ее функциями преобразования являются функции подстановки и перестановки, которые обеспечивают преобразование информации путем ее рассеивания и перемешивания. Изучение и анализ этих функций занимает важное место в теории защиты информации. Они настолько тесно связаны между собой, что порой их трудно разделить.

Подстановки являются наиболее изученными функциями из этой пары. Варианты описания, преобразования и применения подстановок опубликованы довольно широко в научной литературе.

Перестановки являются менее изученными функциями, но по своей значимости они являются не менее важными для теории защиты информации. Обратим внимание на еще одно очень важное обстоятельство. Проведенный анализ, как перестановок, так и подстановок затруднен тем, что в открытых источниках они исследованы и описаны как отдельные объекты, структуры и элементы и не рассмотрены в комплексе, когда одновременному анализу подвергаются все возможные перестановки отдельного множества или даже ряда множеств.

Решение данной задачи может быть основано на методологии научного познания, когда объекты исследования рассматриваются с точки зрения фундаментальных основ и проецируются на известные математические модели и методы. В рамках предложенного подхода комплексное представление о перестановках может дать система счисления ряда факториальных множеств, имеющая (в отличие от известных позиционных систем счисления) переменное основание и переменные позиционные коэффициенты.

Накопленные результаты исследований и анализ полученных материалов показали, что наметилась некоторая неоднозначность в обозначениях и нечеткость ряда определений при проведении исследований свойств таких характеристик подстановок и перестановок как декремент, количество независимых циклов, инверсия, четность и знак подстановок и перестановок, что потребовало проведения более глубокого их теоретического изучения. Одновременно это потребовало в дополнение к качественным

понятиям факториальных множеств, ряда факториальных множеств и его системы счисления разработки аксиоматических основ функций подстановки и перестановки в системе счисления ряда факториальных множеств.

Аксиоматические основы перестановок ряда факториальных множеств, если их рассматривать бессистемно, трудно поддаются общему анализу и проверке на непротиворечивость, полноту и независимость утверждений, поэтому их необходимо упорядочить и структурировать. Наиболее оптимальным вариантом в нашем приложении к группам преобразований (подстановкам и перестановкам), как нам кажется, является способ построения научной теории, при котором в основу теории кладутся некоторые исходные положения, называемые аксиомами, а все остальные предложения теории получаются как логические следствия аксиом. Этот способ называется аксиоматическим методом.

В XIX веке математики обратили свое внимание на дедуктивный способ построения математических теорий, который повлек за собой возникновение связанной с самим понятием аксиоматического метода новой проблематики, на основе которой выросла теория доказательств как основной раздел современной математической логики. Далее возникли проблемы более общего математического характера: непротиворечивость, полнота и независимость той или иной системы аксиом.

Первые результаты в этой области принес метод интерпретаций, который может быть представлен следующим образом. Пусть каждому исходному понятию и отношению данной аксиоматической теории поставлен в соответствие некоторый конкретный математический объект. Совокупность таких объектов называется полем интерпретации. Всякому утверждению теории естественным образом ставится в соответствие некоторое высказывание об элементах поля интерпретации, которое может быть истинным или ложным. Тогда говорят, что утверждение теории соответственно истинно или ложно в данной интерпретации.

Метод интерпретаций позволяет устанавливать факт относительной непротиворечивости. Он также позволяет решать вопрос о независимости аксиом. Для доказательства того, что выбранная аксиома некоторой теории не выводима из остальных аксиом этой теории и, следовательно, необходима для получения всего объема данной теории, достаточно построить такую интерпретацию теоретических знаний, в которой выбранная аксиома была ложной, а все остальные аксиомы этой теории истинны.

Слабая сторона метода интерпретации состоит в том, что в вопросах непротиворечивости и независимости систем аксиом он дает возможность получать только результаты, носящие относительный характер. Важным достижением этого метода стал факт, что с его помощью была выявлена особая роль арифметики как такой математической теории, к вопросу непротиворечивости которой сводится аналогичный вопрос для целого ряда других теорий.

На данном этапе накопления и осмысления информации по аксиоматическим основам подстановок ряда факториальных множеств мы еще не можем построить полную научную теорию, но полученные результаты уже можно некоторым образом систематизировать и структурировать, используя их как базу для дальнейших исследований.

Для реализации поставленной цели возьмем за основу аксиоматическое построение теории, которое можно применять для любой дисциплины, научной теории или некоторых объектов. Аксиоматическое построение включает в себя:

- 1) список основных, неопределяемых понятий (объектов, отношений, структур и т.д.);
- 2) список утверждений, не подлежащих доказательству (аксиом);
- 3) определения других, более сложных понятий;
- 4) формулировки и доказательства других утверждений (теорем, лемм, следствий и т.д.).

Приведем результаты аксиоматического построения для подстановок ряда факториальных множеств. Эти результаты будут расширены и уточнены при рассмотрении операций над подстановками и перестановками (умножение, деление, сложение, вычитание, возведение в степень и обратная к ней операция); классификация подстановок и перестановок, построение и анализ их классов; более глубоком анализе общих характеристик подстановок и перестановок (количество независимых циклов, декремент, инверсия, четность и знак).

На данном этапе аксиоматические основы функций подстановки в системе счисления ряда факториальных множеств как аксиоматическое построение можно представить в следующем виде.

Список основных, неопределяемых понятий. Под неопределяемыми понятиями понимается то, что данные понятия не определяются в данной работе, но они уже определены в других работах, разделах математики или специальных дисциплинах.

Список основных, неопределяемых понятий включает в себя:

- множество;
- множество натуральных чисел, натуральный ряд чисел, ноль;
- десятичная, двоичная и другие позиционные системы счисления;
- правила построения позиционной системы счисления;
- факториал;
- перестановка, подстановка;

Список утверждений, не подлежащих доказательству (аксиом, постулатов) включает в себя:

- аксиомы Пеано (аксиома индукции, аксиома Архимеда, принцип математической индукции, обобщенный принцип математической индукции),
- основные определения (аксиомы) ряда факториальных множеств;
- аксиомы групп и групп преобразований как составная часть аксиом алгебраических структур, симметрических и циклических групп;

Определения других, более сложных понятий включают в себя:

- позиционные системы счисления с постоянным основанием;
- позиционные системы счисления с переменным основанием;
- алгоритмы перевода чисел из одной позиционной системы счисления в другую;
- множества и его характеристики (номер, количество элементов, мощность и т.д.);
- позиционный метод формирования множеств;
- основные характеристики позиционных множеств;
- понятие предыдущего и последующего множеств;
- образующие элементы и множество образующих элементов;
- факториальное множество;
- ряд факториальных множеств;
- принципы и этапы формирования системы счисления ряда факториальных множеств;
- система определений и структура ряда факториальных множеств;
- перестановка и ее внутренняя структура;
- перестановка как группа преобразований;
- система счисления ряда факториальных множеств;
- характеристики подстановок (количество независимых циклов, декремент, инверсия, знак, четность) и результаты их анализа.

Формулировки и доказательства других утверждений (теорем, лемм, следствий).

На данном этапе авторами предложены утверждения, теоремы, леммы и следствия. Их доказательства кажутся авторам очевидными на данном этапе. Одновременно с этим аксиоматика данного направления не является законченной, ее непротиворечивость очевидна, но не доказана. Там не менее полученные результаты не мешают проведению исследований характеристик, свойств и особенностей функций подстановки и

перестановки и их взаимодействию с другими криптографическими функциями и алгебраическими структурами в процессе создания и анализа современных информационных систем.

Предложенные авторами в процессе проведения исследований теоремы можно распределить по следующим направлениям:

- теоремы о числе элементов множеств;
- теоремы о номере, числе образующих элементов и числе подмножеств факториальных множеств;
- таблица распределение теорем по факториальным множествам;
- теоремы о числе элементов и числе образующих элементов перестановки;
- теоремы преобразования чисел позиционных систем счисления и системы счисления ряда факториальных множеств.

В перспективе к данному подразделу можно будет отнести теоремы, связанные с характеристиками подстановок, действиям выполняемыми с подстановками и классами подстановок.

Библиографический список

1. Мартынов А.П., Мартынова И.А., Фомченко В.Н. Аксиоматические основы функций подстановки в системе счисления ряда факториальных множеств и их характеристики. Монография. – Саров: ФГУП «РФЯЦ-ВНИИЭФ», 2019. 210 с.
2. Мартынов А.П., Мартынова И.А. Функции перестановки в системе счисления ряда факториальных множеств. Вестник ВГУ, Серия: Системный анализ и информационные технологии, № 3, 2016, стр. 42-49
3. Свидетельство о государственной регистрации программы для ЭВМ № 2020613795. Программный комплекс анализа подстановок ряда факториальных множеств / Мартынов А.П., Николаев Д.Б., Сплюхин Д.В., Фомченко В.Н., Мартынова И.А. Зарег. 23.03.2020 г. – М.: Роспатент, 2020.
4. Мартынов А.П., Николаев Д.Б., Мартынова И.А. Криптографические системы и метод факториального сжатия информации. – Известия института инженерной физики, 2016. № 4(42). С.54-57.
5. Пат. 2408924, 10.01.2011. Устройство обеспечения безопасного формирования параметров / Курочкин А.А., Мартынов А.П., Николаев Д.Б., Фомченко В.Н. Патент на изобретение № 2408924 от 10.01.2011.
6. Пат. 2623894, 07.09.2017. Способ преобразования данных с равновероятностной инициализацией / Мартынов А.П., Мартынова И.А., Марунин М.В., Николаев Д.Б., Фомченко В.Н. Патент на изобретение № 2623894 от 07.09.2017.
7. Пат. 2546238, 07.10.2013. Способ транслитерационного преобразования информации и передачи ее по каналам связи / Грибунин В.Г., Мартынов А.П., Николаев Д.Б., Фомченко В.Н. Патент на изобретение № 2546238 от 07.10.2013.
8. Пат. 2554525, 07.10. 2013.Способ преобразования информации с синхронной сменой инициализирующих последовательностей в блоках, соединенных каналом связи с неопределенным периодом смены / Фомченко В.Н., Мартынов А.П., Николаев Д.Б., Масыгин А.М., Мартынов А.А. Патент на изобретение № 2554525 от 07.10.2013.
9. Пат. 2256053. Устройство управления электронными замками / Гончаров С.Н., Мартынов А.П., Николаев Д.Б., Фомченко В.Н. Заявл. 20.01.2003, опубл. 10.07.2005. Бюл. № 19.
10. Пат. 2224075. Электронно-механический замок / Мартынов А.П., Николаев Д.Б., Сычев В.Т., Шишкин Г.И. Заявл. 28.11.2001, опубл. 20.02.2004. Бюл. № 5.
11. Пат. 2154722. Электронный кодовый замок / Мартынов А.П., Николаев Д.Б., Шишкин Г.И. Заявл. 27.08.1999, опубл. 20.08.2000. Бюл. № 23.

I.A. Martynova, D.B. Nikolaev

**METHODOLOGY OF SCIENTIFIC KNOWLEDGE ON THE EXAMPLE
OF FORMATION OF AXIOMATIC APPROACH IN THE ELEMENTARY ALGEBRA**

Sarov Institute of Physics and Technology - branch of NRNU MEPhI

The description of the interpretive approach for analytical studies of elements of a number of factorial sets is given, the study of the factors of expansion of the series of factorial sets when using modern mathematical information transformations is carried out.

Keywords: series of factorial set, decrement, inversion, inverse function, multiplication of sets, commutants.

КРИТЕРИИ ОЦЕНКИ КАЧЕСТВА МЕТОДОВ ПРЕОБРАЗОВАНИЯ ДАННЫХ НА ОСНОВЕ ТЕОРИИ МНОЖЕСТВ

Саровский физико-технический институт – филиал НИЯУ МИФИ

Рассмотрены вопросы формирования критериев оценки качества методов обеспечения безопасности информации в рамках построения общей теории системного анализа информационных (криптографических) систем, их моделей и алгоритмов.

Ключевые слова: ключ, криптографическая система, преобразование, системный анализ, стойкость.

Важным аспектом обеспечения безопасности, в информационной системе, в том числе и в образовательной информационной системе, является наличие надежной составляющей защиты данных. В качестве такой составляющей используются криптографические алгоритмы и протоколы преобразования данных, объединенные в криптографические системы. Результаты анализа основополагающих работ по криптографии [1-4] показывают, что криптографическая система в общем случае определяется, как некоторое отображение множества сообщений M на множество криптограмм E представленное как $M \rightarrow E$ (в общем случае эти множества могут быть неравными).

Каждое отображение соответствует способу шифрования при помощи конкретного ключа K_i . Общая модель классической криптографической системы, соответствующая этим отображениям, приведена на рис. 1. Она предполагает наличие таких составляющих ее элементов как источник сообщения и источник ключа, шифратора и дешифратора со своими характеристиками, открытый канал связи для передачи криптограмм, доступный злоумышленнику и защищенный канал для передачи ключа. Составные элементы модели могут анализироваться и синтезироваться как совместно, так и отдельно.

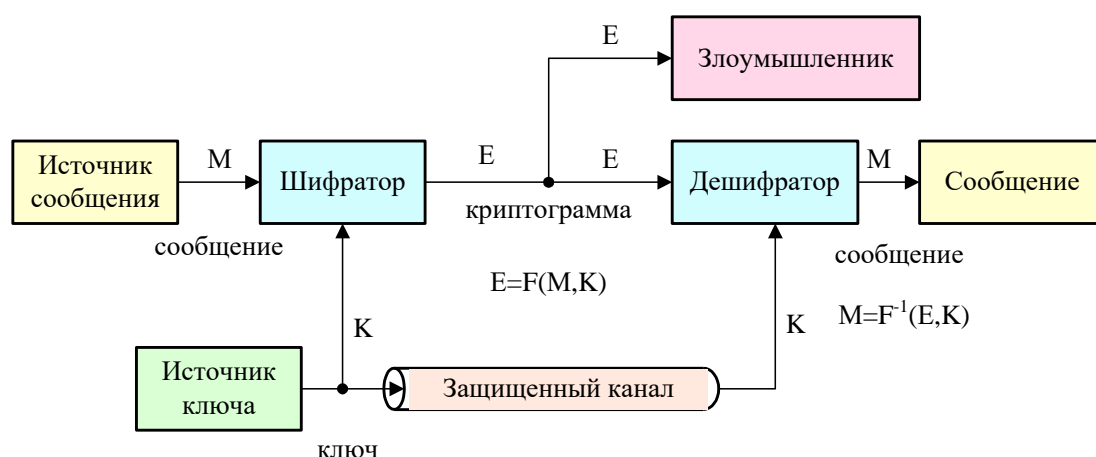


Рис. 1. Общая модель классической криптографической системы

В таком виде классическая криптографическая система рассматривается в большинстве работ по информационным и криптографическим системам. Одновременно с этим в последнее время получила развитие теория систем и системного анализа, предоставляющая относительно развитый инструментарий для моделирования и анализа систем и было бы целесообразным применение их достижения к анализу криптографических систем.

С точки зрения теории систем и системного анализа, изучающих общие положения, законы, принципы построения, функционирования, анализа, а также основ моделирования структура криптографической системы состоит из взаимосвязанных элементов и групп элементов, объединенных связями. Каждый элемент системы можно рассматривать как отдельную подсистему и разрабатывать и анализировать как совместно, так и независимо друг от друга. Связь между элементами обеспечивает возникновение и сохранение структуры и свойств криптографической системы. Среда, в которой функционирует анализируемая система предполагает наличие открытого и закрытого каналов, а также злоумышленника.

Как и в теории систем и системного анализа для моделей криптографических систем объектом изучения для нас являются системы процессов, происходящих внутри них и явлений окружающей действительности, предметом исследования являются общие законы, закономерности, принципы, технологии функционирования и правила проведения системного анализа. Как и любая другая система, она функционирует в системе ограничений, предполагающих некоторые условия и связи, сужающие область функционирования, а также совокупность условий, характеризующих режимы работы, которые являются параметрами системы.

Для практического применения криптографической системы (рис.1) выбирают ключ и посылают его в точку приема по защищенному каналу [1-3]. Затем выбирают или формируют сообщение и с помощью преобразования, соответствующего выбранному ключу, получают криптограмму, которая является функцией ключа и сообщения

$$E = F(M, K). \quad (1)$$

Криптограмма передается в точку приема, в которой выполняют обратные преобразования и восстанавливают первоначальное сообщение

$$M = F^{-1}(E, K). \quad (2)$$

Источник ключа генерирует конечное число ключей, каждому из которых соответствует некоторая вероятность. При анализе ключевых последовательностей необходимо обращать внимание на объем ключа и его характер (постоянный, случайный или псевдослучайный) и частоту смены (ключи сеансов).

Рассмотрим шифратор и дешифратор которые являются преобразующими элементами криптографической системы. Они, как и вся система в целом, в общем случае могут быть представлены в виде схемы простейшей системы, приведенной на рис. 2.

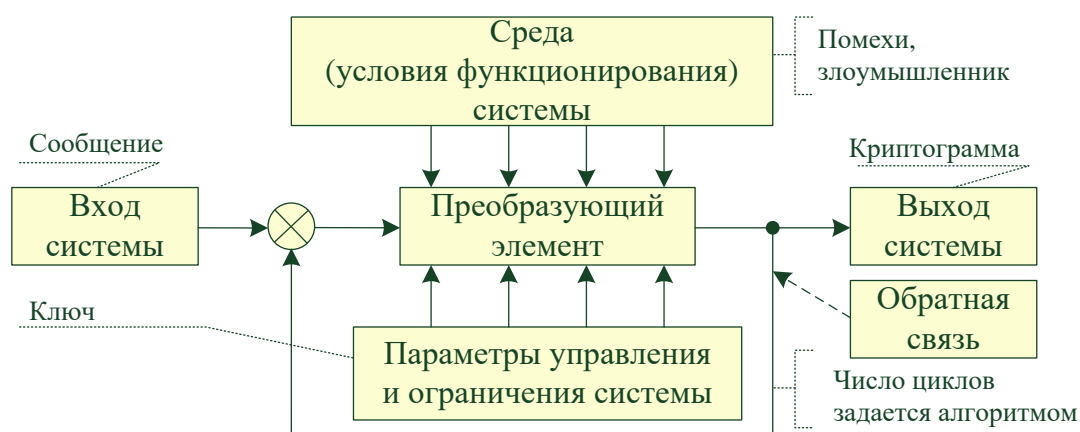


Рис. 2. Элемент криптографической системы, соответствующий схеме простейшей системы

В канале связи присутствуют помехи, которые могут быть как естественными, так и искусственно создаваемыми [4]. К каналу связи имеет доступ злоумышленник. Если злоумышленник перехватит криптограмму, то у него появляется возможность вычислить исходное сообщение и/или ключ.

Иногда удобнее понимать криптограмму не как функцию двух переменных, а как некоторое семейство операций или отображений элементов множеств, записанных в виде

$$E = T_i M, \quad (3)$$

которое означает, что отображение T_i (соответствующее некоторому ключу) примененное к сообщению M , дает криптограмму E [1-4].

В криптографической системе должна иметься возможность восстанавливать исходное сообщение M на приемном конце, когда известны криптограмма E и ключ K , поэтому отображение T_i должно иметь единственное обратное отображение T_i^{-1}

$$T_i T_i^{-1} = I, \quad (4)$$

где I – тождественное отображение. Таким образом

$$M = T_i^{-1} E. \quad (5)$$

Это обратное отображение T_i^{-1} должно существовать и быть единственным для каждой криптограммы. Именно это направление, рассматриваемое с точки зрения теории множеств, является наиболее перспективным для дальнейших исследований, так как позволяет анализировать систему в общем виде с применением математической теории групп колец и полей [5-7].

Состояние элементов системы при анализе может быть статическим, динамическим и переходным, оно предполагает наличие ряда векторов [9, 10]:

- $\vec{U}_\phi = \vec{U}_\phi(U_1, U_2, U_3, \dots, U_n)$ – вектор фактического состояния;

- $\vec{U}_ж = \vec{U}_ж(U_{ж1}, U_{ж2}, U_{ж3}, \dots, U_{жn})$ – вектор желаемого состояния;

- $\vec{U}_{ус} = \vec{U}_{ус}(U_{ус1}, U_{ус2}, U_{ус3}, \dots, U_{усn})$ – вектор условий;

- $\vec{U}_{вус} = \vec{U}_{вус}(U_{вус1}, U_{вус2}, U_{вус3}, \dots, U_{вусl})$ – вектор управления системой.

Ограничения системы также могут характеризоваться некоторой совокупностью параметров

- $\vec{U}_{огр} = \vec{U}_{огр}(U_{огр1}, U_{огр2}, U_{огр3}, \dots, U_{огрl})$ – вектор ограничений.

В процессе формирования и системного анализа вектора ограничений при проектировании и эксплуатации реальных криптографических систем необходимо учитывать, что существует три вида нападения на систему: прослушивание, подмена и иницирование криптограммы, когда истинной передачи зашифрованного сообщения не было [8-14]. Их оценка производится по степени секретности (стойкости), размеру ключа, сложности операций преобразования (шифрования, дешифрования, кодирования и декодирования), а также по распространению ошибок и изменению объема сообщения. Даже эти краткие замечания и фрагменты показывают важность анализа криптографических систем с точки зрения теории систем и системного анализа.

В результате анализа общей модели классической криптографической системы и элементов теории систем можно сделать вывод, что появилась необходимость построения общей теории системного анализа информационных и криптографических систем, их моделей и алгоритмов. Это позволит объединить достигнутые результаты теории систем и системного анализа с достижениями в области информационных и криптографических систем.

Библиографический список

1. Шеннон К. Работы по теории информации и кибернетике. Теория связи в секретных системах. М.: ил, 1963.
2. Алферов А.П., Зубов А.Ю., Кузьмин А.С., Черемушкин А.В. Основы криптографии: Учебное пособие. –М.: Гелиос АРН, 2001. –480с., ил.
3. Мартынов А.П., Фомченко В.Н. Криптография и электроника / Под ред. А.И. Астайкина. Саров: ФГУП «РФЯЦ-ВНИИЭФ», 2006, 452 с.
4. Шнайер Б.. Прикладная криптография. Протоколы, алгоритмы, исходные тексты на языке Си. –М.: Издательство ТРУМФ, 2003– 816 с.: ил.

5. Мартынова И.А., Машин И.Г., Фомченко В.Н. Введение в теорию поля и ее приложения: Монография. – Саров: ФГУП «РФЯЦ-ВНИИЭФ», 2014. – 108 с.: ил.
6. Мартынова И.А., Машин И.Г., Фомченко В.Н. Теория поля и защита информации: Монография. - Саров: ФГУП «РФЯЦ-ВНИИЭФ», 2017. – 209 с.: ил.
7. Мартынов А.П., Мартынова И.А., Фомченко В.Н. Аксиоматические основы функций подстановки в системе счисления ряда факториальных множеств и их характеристики: Монография. – Саров: ФГУП «РФЯЦ-ВНИИЭФ», 2019. – 210 с.: ил.
8. Кривошеев О.В., Мартынова И.А., Николаева И.А., Фомченко В.Н. Защита результатов интеллектуальной деятельности при разработке модульного программно-технического обеспечения. Сборник докладов 16-ой научно-технической конференции «Молодежь в науке», г. Саров, 2017г.
9. Силич М.П., Силич В.А. Основы теории систем и системного анализа: учеб. пособие / М.П. Силич, В.А. Силич. – Томск: Изд-во Томск. гос. ун-та систем управления и радиоэлектроники, 2013. – 340 с.
10. Вдовин В.М., Суркова Л.Е., Валентинов В.А. Теория систем и системный анализ: Учебник/ В.М. Вдовин, Л.Е. Суркова, В.А. Валентинов. — М.: Издательско-торговая корпорация «Дашков и К°», 2010. — 640 с.
11. Мартынов А.П., Николаев Д.Б., Мартынова И.А. Криптографические системы и метод факториального сжатия информации. – Известия института инженерной физики, 2016. № 4(42). С.54-57.
12. Пат. 2256053. Устройство управления электронными замками / Гончаров С.Н., Мартынов А.П., Николаев Д.Б., Фомченко В.Н. Заявл. 20.01.2003, опубл. 10.07.2005. Бюл. № 19.
13. Пат. 2224075. Электронно-механический замок / Мартынов А.П., Николаев Д.Б., Сычев В.Т., Шишкин Г.И. Заявл. 28.11.2001, опубл. 20.02.2004. Бюл. № 5.
14. Пат. 2154722. Электронный кодовый замок / Мартынов А.П., Николаев Д.Б., Шишкин Г.И. Заявл. 27.08.1999, опубл. 20.08.2000. Бюл. № 23.

I.A. Martynova

CRITERIA FOR ASSESSING THE QUALITY OF DATA CONVERSION METHODS BASED ON THE THEORY OF SETS

Sarov Institute of Physics and Technology - branch of NRNU MEPhI

The issues of the formation of criteria for assessing the quality of information security methods within the framework of constructing a general theory of system analysis of information (cryptographic) systems, their models and algorithms are considered.

Keywords: key, cryptographic system, transformation, system analysis, strength.

НЕОБРАТИМОЕ ПРЕОБРАЗОВАНИЕ С ЗАДАНЫМИ СВОЙСТВАМИ НА ОСНОВЕ СОВРЕМЕННЫХ АЛГОРИТМИЧЕСКИХ РЕШЕНИЙ

Саровский физико-технический институт – филиал НИЯУ МИФИ

Рассмотрено параметризованное семейство Hamsi-подобных хеш-функций. Описан класс линейных преобразований, коэффициент рассеивания которых равен коэффициенту рассеивания линейного преобразования, используемого в Hamsi. Найдены инвариантные подпространства данного класса. Параметризованный алгоритм изменен в соответствии с опубликованными дифференциальными атаками на Hamsi, для улучшения его стойкости.

Ключевые слова: криптографические хеш-функции, инвариантные подпространства, коэффициент рассеивания, параметризация алгоритма, дифференциальные атаки.

Развитие автоматизированных систем управления объектами сложной функциональной структуры, которая может динамически корректироваться и изменяться в процессе своего функционирования, предопределило необходимость совершенствования средств идентификации и аутентификации. Важным фактором в данном направлении является то, что все программно-технические и алгоритмические решения должны базировать на отечественных разработках и соответствовать современному уровню развития технологий и промышленности. Отечественные аппаратные решения, особенно, в части реализации специфических функций и экстремальных условий эксплуатации, ограничены как на уровне вычислительных возможностей, так и в части скорости обработки и объема хранения данных. Единственным выходом из сложившейся ситуации является построение алгоритмически модифицируемых программных платформ, имеющих возможность адаптации под требуемую конфигурацию аппаратных ресурсов. Применительно к процессам идентификации и аутентификации требуется определить базовые критерии, по которым будет осуществляться адаптация процедуры идентификации объекта в каждом конкретном режиме функционирования системы управления. Подобными критериями, в частности, являются длина идентифицирующего признака, вычислительная сложность алгоритма идентификации, скорость его выполнения и объем памяти, требуемый для проведения вычислительных операций. Актуальность данной работы обуславливается необходимостью применения алгоритма хеширования для формирования и проверки идентификационных признаков объектов с возможностью динамического изменения параметра идентификации. Особенностью разработанной алгоритмической идентифицирующей структуры является использование вычислительно необратимого преобразования, которое не позволяет в случае компрометации идентификатора (идентифицирующего признака) объекта подобрать корректную идентифицирующую команду, особенно это важно при использовании отчуждаемых или переконфигурируемых объектов, в том числе военного назначения, с высокой вероятностью компрометации объекта или его утраты.

Международный и отечественный опыт по созданию криптографических хеш-функций заключается в синтезе криптографических алгоритмов, а также создании стандартов хеширования.

Новизна данной работы обуславливается тем, что полностью параметризованного семейства хеш-функций, где длина результирующего хеш-кода может быть произвольной, в данный момент найдено не было. В открытом доступе существуют параметризованные криптографические хеш-функции, где параметрами являются различные характеристики алгоритмов, однако данные хеш-функции дают ограниченный выбор длины результата.

В качестве кандидата на модификацию рассматривались участники конкурса SHA-3. После изучения представленных на конкурс алгоритмов, за основу была взята криптографическая хеш-функция Hamsi. Данный алгоритм незапатентован и удобен для параметризации. Помимо этого, он построен по хорошо изученной структуре Меркле-Дамгора.

Параметризованный алгоритм Hamsi- n (таблица 1) можно представить с помощью композиции следующих преобразований:

- расширение сообщения $E: \{0,1\}^{n/8} \rightarrow \{0,1\}^n$;
- конкатенация сообщений $C: \{0,1\}^n \times \{0,1\}^n \rightarrow \{0,1\}^{2n}$;
- нелинейные преобразования $P, P_f: \{0,1\}^{2n} \rightarrow \{0,1\}^{2n}$;
- усечение сообщения $T: \{0,1\}^{2n} \rightarrow \{0,1\}^n$.

Цепное значение и результат хеш-функции вычисляются следующим образом:

$$h_i = (T \circ P \circ C(E(M_i), h_{i-1})) \oplus h_{i-1}, h_0 = iv_n, 0 < i < l \quad (1)$$

$$h = (T \circ P_f \circ C(E(M_l), h_{l-1})) \oplus h_{l-1} \quad (2)$$

В алгоритме Hamsi- n при версии n используются блоки в $n/8$ бит.

Таблица 1

Длина блоков для различных версий алгоритма	
Версии	Длина блока
Hamsi-2048	256
Hamsi-1024	128
Hamsi-512	64
Hamsi-256	32
Hamsi-128	16

В качестве начальных значений вектора iv_n алгоритма Hamsi- n берутся первые n бит кодировки UTF-8 сообщения "Ozgul Kucuk, Katholieke Universiteit Leuven, Departement Elektrotechniek, Computer Security and Industrial Cryptography, Kasteelpark Arenberg 10, bus 2446, B-3001 Leuven-Heverlee, Belgium.". При длине начального вектора большего, чем длина данного сообщения, сообщение произвольно дополняется ненулевыми значениями.

Последний блок открытого текста дополняется битом 1 и последующими битами 0 до кратности длины блока, после чего добавляется еще один или несколько блоков, отображающих длину в битах открытого текста до дополнения. Для расширения сообщения используются линейные коды. Блок сообщения представляется в виде вектора, который, впоследствии, умножается на генерирующую матрицу. Для версии алгоритма Hamsi- n размер порождающей матрицы равен $\frac{n}{16} \times \frac{n}{2}$, элементы которой принадлежат F_4 .

При выборе параметров необходимо выбрать такой наилучший линейный код, при котором минимальное количество отличающихся бит после расширения будет максимально. Поэтому вместо использования поля F_4 , можно использоваться поля F_2 или F_{16} .

Тогда размер порождающей матрицы может быть равен:

- размером $\frac{n}{8} \times n$, элементы которой принадлежат F_2 ;
- размером $\frac{n}{16} \times \frac{n}{2}$, элементы которой принадлежат F_4 ;
- размером $\frac{n}{32} \times \frac{n}{4}$, элементы которой принадлежат F_{16} .

Для нахождения наилучшего линейного кода можно воспользоваться существующими базами наилучших линейных кодов, например, при помощи пакетов Magma. Из найденных линейных кодов необходимо определить, при каком линейном коде количество задействованных s -блоков будет наибольшим при изменении 1 бита сообщения. Также существует вероятность, что в базах линейных кодов отсутствует

искомый наилучший линейный код, в таком случае можно воспользоваться полем F_2 , так как оно является наиболее исследованным. В таблице 2 представлены примеры наилучших линейных кодов для Hamsi-128 и Hamsi-256. Так как в базах, на данный момент, не содержится известных наилучших линейных кодов поля F_{16} , данные коды можно рассматривать позже, когда данная сфера будет более исследована.

Таблица 2

Примеры различных используемых линейных кодов

	Код для Hamsi-128	Код для Hamsi-256
Поле F_2	[128,16,52]	[256,32,96]
Поле F_4	[64,8,43]	[128,16,70]

Пусть $m_i - \frac{n}{8}$ битная часть расширенного сообщения, $m_i - \frac{n}{8}$ битная часть итеративного значения, $s_i - \frac{n}{8}$ битная часть внутреннего состояния. Для версии алгоритма Hamsi- n используется версия конкатенации Hamsi-256:

$$C(m_0, m_1, \dots, m_7, c_0, c_1, \dots, c_7) := (m_0, m_1, c_0, c_1, c_2, c_3, m_2, m_3, m_4, m_5, c_4, c_5, c_6, c_7, m_6, m_7), m_i, c_i \in F_2^{\frac{n}{8}}. \quad (3)$$

Результат операции конкатенации можно увидеть на рис. 1.

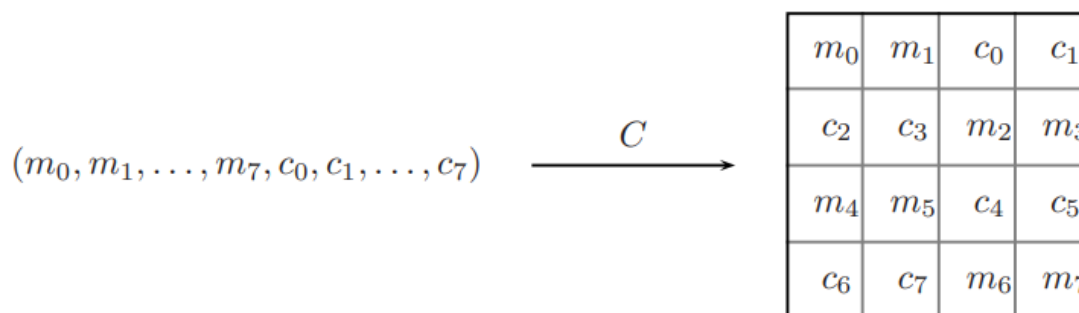


Рис. 1. Графическое представление преобразования конкатенации

Нелинейное преобразование состоит из 3 слоев:

1. Прибавление константных значений и счетчика.
2. Прохождения слоя из S-блоков.
3. Использование линейного преобразования.

Длина констант, используемых в данном преобразовании, равна длине выбранных блоков ($\frac{n}{8}$ бит). В качестве констант могут быть взяты части констант алгоритма Hamsi-512 или же собственные ненулевые, не совпадающие между собой значения. Помимо прибавления константных значений, первый элемент внутреннего состояния s_0 складывается со счетчиком, где счетчик равен номеру раунда преобразования.

В преобразовании через слой S-блоков (таблица 3) проходит все 16 элементов матрицы внутреннего состояния. S-блок идентичен таковому из алгоритма Hamsi-256. На каждый S-блок поступают столбцы матрицы внутреннего состояния.

Таблица 3

Используемый S-блок

x	0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
$S[x]$	8	6	7	9	3	12	10	15	13	1	14	4	0	11	5	2

Рассеивающий слой основан на нескольких применениях линейного преобразования L . Преобразование принимает и преобразует 4 входных слова по $\frac{n}{8}$ бит. Преобразование применяется следующим образом:

$$(s_0, s_5, s_{10}, s_{15}) := L(s_0, s_5, s_{10}, s_{15}), \quad (4)$$

$$(s_1, s_6, s_{11}, s_{12}) := L(s_1, s_6, s_{11}, s_{12}), \quad (5)$$

$$(s_2, s_7, s_8, s_{13}) := L(s_2, s_7, s_8, s_{13}), \quad (6)$$

$$(s_3, s_4, s_9, s_{14}) := L(s_3, s_4, s_9, s_{14}). \quad (7)$$

Для версии Hamsi- n существует необходимость использования другого линейного преобразования, отличающегося от такового в Hamsi-256. Для этого, в качестве линейного преобразования берется преобразование из класса линейных преобразований. Преобразование будет иметь вид:

$$x_1 := x_1 \oplus x_2 \oplus x_3, \quad (8)$$

$$x_4 := x_4 \oplus x_2 \oplus (x_3 \ll c_1), \quad (9)$$

$$x_1 := (x_1 \lll c_2), \quad (10)$$

$$x_4 := (x_4 \lll c_3), \quad (11)$$

$$x_2 := x_2 \oplus x_1 \oplus x_4, \quad (12)$$

$$x_3 := x_3 \oplus x_1 \oplus (x_4 \ll c_4), \quad (13)$$

где $x_1, x_2, x_3, x_4 \in F_2^{\frac{n}{8}}, c_i \in \{1, \dots, \frac{n}{8} - 1\}$.

Данный класс образован различными биективными отображениями $f: A \rightarrow B, A = \{a, b, c, d\}, B = \{x_1, x_2, x_3, x_4\}$, (14)

$$f: A \rightarrow B, A = \{a, b, c, d\}, B = \{x_1, x_2, x_3, x_4\},$$

значениями сдвигов c_1, \dots, c_4 , а также всевозможными комбинациями циклических сдвигов частей вектора (a, b, c, d) на различные значения. При использовании данного линейного преобразования, необходимо использовать такие значения c_1, \dots, c_4 , что $\text{НОД}(c_1, \dots, c_4) = 1$. Это делается для того, чтобы для полученного преобразования не существовало инвариантных подпространств.

Слой усечения $T: \{0,1\}^{2 \cdot n} \rightarrow \{0,1\}^n$ для алгоритма Hamsi- n выглядит следующим образом

$$T(s_0, s_1, \dots, s_{14}, s_{15}) := (s_0, s_1, s_2, s_3, s_8, s_9, s_{10}, s_{11}), s_i \in F_2^{\frac{n}{8}}. \quad (15)$$

В качестве взятия константных значений для преобразования P_f берутся значения по аналогии с преобразованием P .

Число раундов в алгоритме Hamsi- n при $n \leq 256$ равно 3 для преобразования P и 6 для P_f . При $n > 256$ и каждом увеличении версии число раундов увеличивается в два раза для P и P_f соответственно. Примеры числа раундов представлены в таблице 4. Стоит отметить, что число раундов является настраиваемым параметром, и может быть изменено в меньшую сторону для увеличения производительности или в большую для улучшения стойкости алгоритма.

Таблица 4

Число раундов для различных версий алгоритма

Версии	Число раундов P	Число раундов P_f
Hamsi-2048	24	48
Hamsi-1024	12	24
Hamsi-512	6	12
Hamsi-256	3	6
Hamsi-128	3	6

В алгоритме Hamsi- n при значениях версий $n < 256$ и использовании блоков длиной в $\frac{n}{8}$ бит, длина блоков будет принимать относительно малые значения, из-за чего полный перебор будет занимать короткий промежуток времени. Из-за этого рекомендуется использовать версию Hamsi-256 или Hamsi-512 параметризованного алгоритма, улучшающие стойкость за счет увеличения количества производимых

машинных операций при переборе значений, в связи с чем увеличится количество времени, необходимое на атаку алгоритма.

В качестве результата хеш-функции берется конкатенация первых $\frac{n}{8}$ бит каждого блока цепного значения после слоя усечения последнего раунда.

С целью улучшения стойкости параметризованной версии были проанализированы существующие работы по криптоанализу алгоритма Hamsi. Основными опубликованными работами являются дифференциальные атаки на функцию сжатия алгоритма Hamsi-256, которые не понижают стойкость самой хеш-функции. После анализа данных атак были внесены следующие изменения в функцию сжатия версии Hamsi- n :

1. Изменен используемый S-блок.
2. Внесены изменения в полученный класс линейных преобразований.
3. Увеличено количество применений линейного преобразования к внутреннему состоянию на одном раунде.

После внесения данных изменений был проведен анализ того, являются ли успешными данные дифференциальные атаки на функцию сжатия. Был сделан вывод, что внесенные изменения увеличили стойкость функции сжатия к дифференциальным атакам такого вида.

Также в работе описаны способы эффективной реализации данного алгоритма. Для оценки скорости работы различных версий Hamsi- n было подсчитано среднее число битовых операций на выполнение одной функции сжатия:

$$n \cdot \left(23 \cdot x + 1 + \frac{n}{8 \cdot p} \right),$$

где n - версия алгоритма, x - число раундов, p - число бит которое обрабатывается за одно обращение к таблице поиска. Сделан вывод, что при сравнении версий Hamsi- n и Hamsi- $2 \cdot n$, для обработки информации одинакового объема для версии Hamsi- n требуется примерно в 2 раза меньше битовых операций. Помимо этого, было проведено сравнение 256 и 512 битных версий алгоритма с участниками конкурса SHA-3.

Поставленная цель, заключающаяся в синтезе параметризованной криптографической хеш-функции была достигнута. В результате проведенного анализа был получен алгоритм Hamsi- n вырабатывающий n бит результирующего хеш-кода.

Библиографический список

1. Мартынов А.П., Мартынова И.А., Фомченко В.Н. Аксиоматические основы функций подстановки в системе счисления ряда факториальных множеств и их характеристики. Монография. – Саров: ФГУП «РФЯЦ-ВНИИЭФ», 2019. 210 с.
2. Мартынов А.П., Мартынова И.А. Функции перестановки в системе счисления ряда факториальных множеств. Вестник ВГУ, Серия: Системный анализ и информационные технологии, № 3, 2016, стр. 42-49
3. Свидетельство о государственной регистрации программы для ЭВМ № 2020613795. Программный комплекс анализа подстановок ряда факториальных множеств / Мартынов А.П., Николаев Д.Б., Сплюхин Д.В., Фомченко В.Н., Мартынова И.А. Зарег. 23.03.2020 г. – М.: Роспатент, 2020
4. Мартынов А.П., Николаев Д.Б., Мартынова И.А. Криптографические системы и метод факториального сжатия информации. – Известия института инженерной физики, 2016. № 4(42). С.54-57.
5. Пат. 2408924, 10.01.2011. Устройство обеспечения безопасного формирования параметров / Курочкин А.А., Мартынов А.П., Николаев Д.Б., Фомченко В.Н. Патент на изобретение № 2408924 от 10.01.2011.

6. Пат. 2623894, 07.09.2017. Способ преобразования данных с равновероятностной инициализацией / Мартынов А.П., Мартынова И.А., Марунин М.В., Николаев Д.Б., Фомченко В.Н. Патент на изобретение № 2623894 от 07.09.2017.
7. Пат. 2546238, 07.10.2013. Способ транслитерационного преобразования информации и передачи ее по каналам связи / Грибунин В.Г., Мартынов А.П., Николаев Д.Б., Фомченко В.Н. Патент на изобретение № 2546238 от 07.10.2013.
8. Пат. 2554525, 07.10. 2013.Способ преобразования информации с синхронной сменой инициализирующих последовательностей в блоках, соединенных каналом связи с неопределенным периодом смены / Фомченко В.Н., Мартынов А.П., Николаев Д.Б., Масягин А.М., Мартынов А.А. Патент на изобретение № 2554525 от 07.10.2013.
9. Пат. 2256053. Устройство управления электронными замками / Гончаров С.Н., Мартынов А.П., Николаев Д.Б., Фомченко В.Н. Заявл. 20.01.2003, опубл. 10.07.2005. Бюл. № 19.
10. Пат. 2224075. Электронно-механический замок / Мартынов А.П., Николаев Д.Б., Сычев В.Т., Шишкин Г.И. Заявл. 28.11.2001, опубл. 20.02.2004. Бюл. № 5.
11. Пат. 2154722. Электронный кодовый замок / Мартынов А.П., Николаев Д.Б., Шишкин Г.И. Заявл. 27.08.1999, опубл. 20.08.2000. Бюл. № 23.

K.D. Ermakov D.V. Splyukhin

INVERSIBLE TRANSFORMATION WITH GIVEN PROPERTIES BASED ON MODERN ALGORITHMIC SOLUTIONS

Sarov Institute of Physics and Technology - branch of NRNU MEPhI

A parameterized family of Hamsi-like hash functions is considered. A class of linear transformations is described, the dissipation factor of which is equal to the dissipation factor of the linear transform used in Hamsi. Invariant subspaces of this class are found. The parameterized algorithm has been modified in accordance with the published differential attacks on Hamsi to improve its robustness.

Keywords: hash functions, invariant subspaces, scattering coefficient, algorithm parameterization, differential attacks.

**МОДЕЛИ БЕЗОПАСНЫХ ИНФОРМАЦИОННЫХ СИСТЕМ,
ОРИЕНТИРОВАННЫХ НА ДИСТАНЦИОННОЕ ОБРАЗОВАНИЕ***Саровский физико-технический институт – филиал НИЯУ МИФИ*

Представлено решение задачи создания безопасной информационной системы за счет разработки стойкого криптографического преобразования.

Ключевые слова: криптографическое преобразование, алгоритм шифрования, кодирование, ключ, криптографическая система.

Введение ограничений на проведение очных аудиторных занятий привело к необходимости стремительного роста систем дистанционного обучения, в рамках которых требуется создание безопасных каналов взаимодействия между преподавателем и аудиторией в индивидуальном или групповом формате. Разглашение элементов информационного взаимодействия может привести к возникновению конфликтных ситуаций как при проведении контроля и оценки знаний, так и при утере персональных данных взаимодействующих субъектов. Решение данной проблемы лежит в применении криптографических алгоритмов, которые позволяют обеспечить защиту передаваемых данных на требуемом уровне. В основе любого способа криптографического преобразования лежит алгоритм шифрования, предназначенный для решения задачи обеспечения конфиденциальности информации. Шифрование определяется как взаимнообратное преобразование незащищенной информации (открытой) информации в зашифрованную (закрытую) форму – шифртекст, в которой она является недоступной для злоумышленника [1]. При этом используются специальные ключи, позволяющие зашифровывать и (или) расшифровывать информацию. Если для зашифрования и расшифрования используется один и тот же ключ, такая система называется симметричной.

При выборе криптографического алгоритма для использования в конкретной разработке одним из определяющих факторов является его стойкость, т. е. устойчивость к попыткам злоумышленника его раскрыть. Алгоритм шифрования считается криптостойким, если, имея закрытые данные и не зная секретный ключ, невозможно получить информацию об открытых данных. Общепринятая схема построения симметричных криптосистем представляет собой циклические перестановки и подстановки битов в блоке фиксированной длины, алгоритм использования которых определяется секретным ключом. В асимметричных криптосистемах для зашифрования и расшифрования применяют различные, не связанные друг с другом ключи, при этом ключ расшифрования является секретным, а ключ зашифрования – открытым. Этим достигается возможность передавать конфиденциальные сообщения по незащищенному каналу без предварительной передачи секретного ключа.

Для создания стойкой криптографической системы необходимо решить ряд сложных задач, удовлетворяющих определенным условиям. Рассмотрим пример разработки способа криптографического преобразования.

В представленном способе исходная информация на передающей стороне преобразуется с помощью ключевой последовательности, а на приемной стороне информация восстанавливается с помощью той же ключевой последовательности (за основу взят способ [2]), при этом исходные данные обрабатываются посимвольно, входные данные подвергаются преобразованию методом арифметического кодирования с введенными криптографическими преобразованиями, зависящими от ключевой последовательности, и при кодировании преобразуются с помощью кодирующей функции

Е (в качестве функции Е выбирается функция арифметического кодирования согласно [4]), совместно с криптографическими преобразованиями F_1 , F_2 и F_3 в вещественное число из полуинтервала $[0,1)$, которое записывается двоичной последовательностью конечной переменной длины и вместе с ключевой последовательностью однозначно определяет входные данные.

При этом в качестве функции преобразования F_2 используем функцию линейного конгруэнтного преобразования вида $F(x)=ax+b \bmod c$, с начальным значением x , зависящим от ключевой последовательности, а в качестве составляющей функций преобразования F_1 и F_3 используем нормированную функцию линейного конгруэнтного преобразования вида

$$F(x) = \frac{ax+b \bmod c}{N}, \quad (1)$$

с начальным значением x (т.е. x представляет собой последовательность значений ключевой последовательности), зависящим от ключевой последовательности и разным для обеих F_1 и F_3 , где константы, a и b выбираются так, чтобы вычисленная последовательность не обладала периодом максимальной длины, а константа c для F_1 и F_3 выбирается равной целой части $0.25N$, где N - фиксированное простое целое число большого порядка, выбранное произвольно до начала кодирования.

Другими словами, положительный результат достигается по существу за счет введения посимвольных криптографических преобразований, зависящих от данной ключевой последовательности, с использованием метода арифметического кодирования двоичной информации, который состоит в том, что при кодировании входные данные преобразуются с помощью кодирующей функции Е в вещественное число из полуинтервала $[0,1)$, который называют отображающим интервалом. Результат функции Е [3, 4] является вещественным числом из отображающего интервала, которое записывается двоичной последовательностью конечной длины и однозначно определяет входные данные. Таким образом, входные и выходные данные представляют собой двоичную последовательность переменной длины. Перед началом процесса кодирования входные данные разбиваются на блоки фиксированной длины, и кодируются поочередно. Значением блока является двоичное представление битов блока. Совокупность значений всех блоков образует алфавит, а сами значения блоков являются символами алфавита. Количество возможных значений формируемых блоков из входной последовательности данных обозначается через M . По мере кодирования входных данных, отображающий интервал уменьшается, а количество бит для его представления возрастает.

Каждому символу алфавита ставят в соответствие априорную вероятность его появления в потоке входных данных. Дополнительно каждому символу присваивают интервал вероятности на промежутке $[0, 1)$, такой, что его длина равна вероятности появления данного символа [6]. Нижнюю границу вероятности обозначим через L , а верхнюю границу вероятности через H . Совокупность пар $[L_{(\text{символ})}, H_{(\text{символ})})$ для всех символов входных данных однозначно определяет вероятностную модель появления каждого символа в кодируемом тексте.

Очередные кодируемые символы входных данных сокращают величину отображающего интервала исходя из значений их априорных вероятностей, определяемых имеющейся моделью. Более вероятные символы делают это в меньшей степени, чем менее вероятные, добавляя меньше битов к выходным данным. В процессе кодирования вероятностная модель изменяется так, чтобы учитывать фактические появления символов во входных данных. Данный эффект достигается за счет увеличения вероятности появления текущего символа и соответственного уменьшения всех остальных. Таким образом, достигается оптимальное сжатие входных данных [3,4]. Данный метод сжатия называют адаптивным сжатием. Для того чтобы сигнализировать окончание кодирования используют специальный символ, который обозначают EOF.

Начальное построение вероятностной модели и ее изменение при обработке очередного символа должны совпадать для процессов кодирования и декодирования, иначе входные данные не будут должным образом декодированы.

Согласно описываемому способу, начальное построение вероятностной модели, а также ее изменение при обработке очередного символа модифицируется с использованием ключевой последовательности K , называемой также ключом. Таким образом, не имея должного ключа, невозможно адекватно восстановить исходные входные данные, преобразованные с помощью данного способа. Для достижения этого вводятся три дополнительных вида криптографических преобразований в алгоритм кодирования и, в частности, в вероятностную модель символов входных данных.

Первый вид дополнительного криптографического преобразования изменяет начальную таблицу априорных вероятностей увеличением вероятности появления одних символов и уменьшением других, с помощью некоторой функции F_1 в зависимости от ключевой последовательности K . Данный вид дополнительного криптографического преобразования применяется в начале процесса кодирования и декодирования один раз. Второй и третий виды дополнительного криптографического преобразования действуют на каждом шаге кодирования, то есть один раз при кодировании каждого символа.

Второй вид дополнительного криптографического преобразования изменяет таблицу интервалов вероятности всех символов, переставляя интервалы вероятности на отрезке $[0, 1)$ с помощью преобразования F_2 в зависимости от ключевой последовательности K и сквозного номера текущего символа. Данный вид дополнительного криптографического преобразования не влияет на длину выходных данных, а влияет лишь на скорость обработки закодированной последовательности при декодировании и потому является наиболее предпочтительным в случае, когда используются не все из предлагаемых видов дополнительного криптографического преобразования вероятностной модели.

Третий вид дополнительного криптографического преобразования изменяет таблицу интервалов вероятности всех символов, переназначая каждому символу новый интервал вероятности на отрезке $[0, 1)$ с помощью преобразования F_3 в зависимости от ключевой последовательности K и сквозного номера текущего символа. Данный вид дополнительного криптографического преобразования, в отличие от второго и первого видов, влияет на длину выходных данных.

В данном способе в качестве криптографического преобразования данных используются все три вида дополнительных криптографических преобразований алгоритма арифметического кодирования. В качестве преобразования F_1 выбирается псевдослучайная функция, зависящая от ключевой последовательности K , с вещественным результатом на отрезке $[0, 0.25]$. В качестве преобразования F_2 выбирается псевдослучайная функция, зависящая от ключевой последовательности K , с целочисленным результатом на отрезке $[1, M]$. Множество ее $2M$ первых значений разбивается на пары (a_q, b_q) , где $q=1, \dots, M$. Затем выполняется перестановка местами символа a_q с символом b_q в таблице символов. В качестве преобразования F_3 выбирается псевдослучайная функция, зависящая от ключевой последовательности K , с целочисленным результатом на отрезке $[1, M]$. Множество ее $2M$ первых значений разбивается на пары (a_q, b_q) , где $q=1, \dots, M$. Затем выполняется обмен интервалов вероятностей символа a_q и символа b_q в таблице интервалов вероятностей.

Функциональная схема устройства для реализации способа представлена на рис. 1.

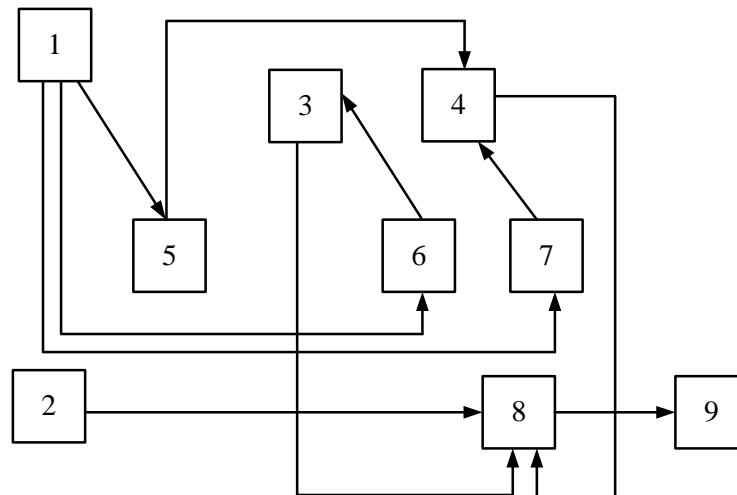
Устройство работает следующим образом. Используя блок 1 (алфавитно-цифровое устройство ввода информации, как, например, описанное в [5]), в устройство вводят секретную ключевую последовательность, которая передается на вход блока 5, блока 6 и блока 7 (операционные блоки, организованные согласно [7]), а выходные данные блока 5 в

блок 4, где преобразовывается таблица вероятностей появления символов. Вероятность P_i каждого i -го ($i=1, 2, 3, \dots, M$) символа увеличивается на результат преобразования F_1 :

$$P_i = P_i + F_1(K, i) \quad (2)$$

После чего все P_i умножаются на нормирующий множитель $0 < \mu \leq 1$, который вычисляется по формуле

$$\mu = \frac{1}{\sum_i P_i} \quad (3)$$



- 1 - блок ввода ключевой последовательности;
- 2 - блок ввода входных данных;
- 3 - блок хранения таблицы обрабатываемых символов;
- 4 - блок хранения таблицы интервалов вероятности для каждого символа;
- 5 - операционный блок, реализующий преобразование F_1 ;
- 6 - операционный блок, реализующий преобразование F_2 ;
- 7 - операционный блок, реализующий преобразование F_3 ;
- 8 - операционный блок, реализующий арифметическое сжатие с использованием блока 2, блока 3 и блока 4;
- 9 - блок выходных данных.

Рис. 1. Устройство для реализации способа

Таким образом, содержимое блока 3 и 4 определяется один раз в начале кодирования с помощью блока 5 и ключевой последовательности. С помощью блока 2 (аналогичного по устройству блоку 1) в устройство по одному символу вводят данные, предназначенные для зашифрования, после чего данные помещаются в блок 8, где посимвольно кодируются с использованием таблиц из блока 3 и блока 4 [3]. Затем на каждом шаге кодирования содержимое блока 3 обновляется с помощью блока 6 следующим образом: из блока 6 извлекается $2M$ первых значений, которые разбиваются на M пар (a_q, b_q) , где $q=1, \dots, M$. Затем с помощью блока 6 выполняется перестановка местами символа a_q с символом b_q в таблице символов блока 3. Также, на каждом шаге кодирования содержимое блока 4 обновляется с помощью блока 7 следующим образом: из блока 7 извлекается $2M$ первых значений, которые разбиваются на M пар (a_q, b_q) , где $q=1, \dots, M$. Затем выполняется переписывание интервалов вероятностей символа a_q и символа b_q в таблице интервалов вероятностей, то есть:

La_q заменяется на Lb_q и Lb_q заменяется на La_q ;

Na_q заменяется на Nb_q и Nb_q заменяется на Na_q .

Затем входные данные, преобразованные описанным выше способом, посимвольно передаются в блок 9 (алфавитно-цифровое устройство вывода), откуда могут быть непосредственно извлечены. Процесс зашифрования завершается после обработки последнего символа входных данных. Зашифрованные данные извлекаются из блока 9.

При обратном криптографическом преобразовании для расшифрования данных используется устройство, идентичное данному. Расшифрование данных происходит в том же порядке [6-11].

Таким образом, в описанном способе повышается криптостойкость, уменьшается вероятность несанкционированного дешифрования зашифрованной информации, а также снижаются затраты на обработку, передачу и хранение зашифрованной информации. Применение рассмотренного способа криптографического преобразования в системах дистанционного обучения позволяет комплексным образом решить проблему безопасности, целостности и избирательной доступности данных, что является неотъемлемой частью любой распределенной информационной системы, и особенно информационной образовательной системы.

Библиографический список

1. Николаев Д.Б., Мартынов А.П., Фомченко В.Н. Технические средства и методы обеспечения безопасности информации: Учебное пособие. Саров: ФГУП «РФЯЦ-ВНИИЭФ», 2015. – 394с.: ил.
2. Пат. 2130641 С1, кл. G 06 F 13/00, G 09 C 1/00, H 04 L 9/00.
3. Witten Ian H., Neal Radford M., Cleary John G. Arithmetic coding for data compression. Communications of the ACM. - June 1987, Vol.30, №6.
4. Мастрюков Д. Алгоритмы сжатия информации. Арифметическое кодирование. - М.: Монитор, N1, 1994.
5. Клингман Э. Проектирование специализированных микропроцессорных систем. - М.: Мир, 1985.
6. Гмурман В. Теория вероятностей и математическая статистика. - М.: Высшая школа, 2000.
7. Пат. 2497277, 27.10.2013. Способ сжатия двоичных данных в виде структурированных информационных блоков / Николаев Д.Б., Мартынов А.П. Патент на изобретение № 2497277 от 27.10.2013.
8. Астайкин А.И., Мартынов А.П., Николаев Д.Б., Фомченко В.Н. Методы и средства обеспечения программно-аппаратной защиты информации. Научно-техническое издание / Саров, 2015.
9. Пат. 2256053. Устройство управления электронными замками / Гончаров С.Н., Мартынов А.П., Николаев Д.Б., Фомченко В.Н. Заявл. 20.01.2003, опубл. 10.07.2005. Бюл. № 19.
10. Пат. 2224075. Электронно-механический замок / Мартынов А.П., Николаев Д.Б., Сычев В.Т., Шишкин Г.И. Заявл. 28.11.2001, опубл. 20.02.2004. Бюл. № 5.
Пат. 2154722. Электронный кодовый замок / Мартынов А.П., Николаев Д.Б., Шишкин Г.И. Заявл. 27.08.1999, опубл. 20.08.2000. Бюл. № 23.

M.V. Golikhin, K.A. Selivokhin

MODELS OF SAFE INFORMATION SYSTEMS FOCUSED ON DISTANCE EDUCATION

Sarov Institute of Physics and Technology - branch of NRNU MEPhI

The solution to the problem of creating a cryptographic system by developing a strong cryptographic transformation is presented.

Keywords: cryptographic conversion, encryption algorithm, coding, key, cryptographic system.

ПРОГРАММНЫЙ КОМПЛЕКС ДЛЯ АВТОМАТИЗАЦИИ ПРОЦЕССА ДИСТАНЦИОННОГО ТЕСТИРОВАНИЯ

Саровский физико-технический институт – филиал НИЯУ МИФИ

Рассмотрены варианты разграничения доступа к элементам информационного ресурса и обеспечения бесконфликтной работы всех задействованных элементов при автоматизации процессов дистанционного тестирования.

Ключевые слова: запрос-ответ, идентификационные данные, пароль, преобразование, разграничение доступа, уникальные особенности.

Глобальное развитие информационных технологий предопределило полное или частичное внедрение автоматизированных систем во все сферы деятельности человека, будь то крупная промышленность, торговля или образование. Степень автоматизации определяется материальными возможностями, состоянием объекта, алгоритмом функционирования и т.д. В большинстве случаев автоматизированная система представляет собой систему управления и контроля, которая интегрируется в объект [1-3].

Управление и контроль объекта осуществляется при помощи множества специализированных устройств разного назначения, в случае дистанционных образовательных систем – это индивидуальные или групповые автоматизированные рабочие места (АРМ) обучающихся и преподавателей. Изменение режимов работы АРМ, сбор информации, контроль функционирования осуществляется с применением серверных систем различного типа, обеспечивающих поддержание информационных образовательных процессов в работоспособном состоянии [4,8-10]. Индивидуальная работа с каждым обучаемым или группой обучаемых может осуществляться и с вспомогательных АРМ с обязательной передачей всей модифицированной информации на основную серверную систему для исключения параллелизма учебного процесса. Следует отметить, что в случае проведения групповой аттестации обучаемого возникает необходимость работы с АРМ нескольких преподавателей, представляющих собой независимые контролирующие структуры [11,12].

Контролирующие структуры взаимодействуют с объектом по нескольким направлениям и могут осуществлять работу с несколькими группами обучаемых. При этом различные структуры могут работать по одним и тем же направлениям, но с использованием различных параметров. Следует отметить, что данные, используемые разными структурами, различны и в общем случае могут являться конфиденциальной информацией. Непреднамеренный или умышленный доступ к «чужой» информации может повлечь серьезные негативные последствия. В данной ситуации необходимо разграничить доступ к элементам объекта и обеспечить бесконфликтную работу всех задействованных структур [5,13-16]. Решение этой задачи возложено на систему управления и контроля. Рассмотрим варианты взаимодействия различных структур при получении доступа к элементам объекта. Возможны два варианта работы в системе управления:

- все подключаются к элементам объекта с основного АРМ;
- подразделения могут подключаться как с основного АРМ, так и с дополнительных АРМ.

В первом случае необходимо четко идентифицировать представителей преподавательского состава, определять направления их работы и группу обучаемых, с которыми возможно их взаимодействие [17,18].

Во втором случае при подключении вспомогательного АРМ необходимо определить его параметры (идентификационный номер, наличие и сроки разрешения работы с элементами объекта и т.д.) и при необходимости безопасного обмена информации с системой управления выработать необходимые ключевые данные. Наиболее оптимальной является ситуация, в которой связь вспомогательного АРМ и серверной системы осуществляется автоматически без участия оператора [19-22]. В дальнейшем работа осуществляется также, как и в первом случае.

Для идентификации оператора необходимо присвоить ему индивидуальные (уникальные) особенности. Существует большое количество способов идентификации, начиная от простого ввода цифробуквенной информации и заканчивая комплексной оценкой характеристик объекта (пульс, давление, биоритмы и т.д.).

Наибольший эффект наблюдается при комплексном использовании рассмотренных методов, например, объединением паролирования и биометрических методов [26,27]. Однако для применения биометрических методов необходимо сложное и дорогое оборудование, требующее постоянного обновления и корректировки баз данных параметров идентифицируемых объектов. Не менее эффективным на наш взгляд является применение усовершенствованной системы паролирования, когда одна часть пароля хранится у пользователя, а другая находится на специализированном носителе [23,24].

Как уже отмечалось выше при работе с объектом нескольких независимых структур необходимо при активизации очередного сеанса управления перенастраивать некоторые характеристики системы, подключать или отсоединять приборы и устройства и т.д. Таким образом, всю эту информацию необходимо хранить, обеспечивая ее безопасность. Современные операционные системы типа Windows 2000, Windows XP, Linux и т.д. позволяют разграничивать адресное пространство разных пользователей, однако это не гарантирует сохранности информации в случае применения специальных программ взлома, «тройных коней», закладок и т.д. В сложившейся ситуации требуемую информацию целесообразнее хранить на внешнем специализированном носителе информации, например, на носителе вместе с паролем. В качестве подобного носителя может использоваться компакт-диск или интеллектуальное устройство хранения данных [31]. Оператор (пользователь) вводит свою часть пароля с клавиатуры, другая часть считывается со специализированного носителя. Обе части преобразуются определенным образом и формируется результат о разрешении или отказе в доступе [25].

На специализированных носителях храниться пароль и информация, требуемая в процессе работы специалистов определенного подразделения. Все подразделения имеют собственные носители с уникальной касающейся только них информацией о составе подключаемой аппаратуры, драйверами, параметрами технологических процессов, кодами и т.д. При использовании специализированных носителей типа компакт-дисков необходимо обеспечивать защиту хранимых на нем данных и пароля, так как вся информация может быть считана при помощи стандартных приложений операционной системы или в особых случаях специальными программами низкого уровня. Для предотвращения потери данных предлагается применять средства преобразования данных типа алгоритма преобразования данных по ГОСТ 28147-89 или хотя бы стандартные программы, такие как программы сжатия данных с паролем (ARJ, ZIP, RAR и т.д.), т.е. вся информация должна храниться в преобразованном виде. С целью исключения похищения пароля и использования его с другим массивом данных в качестве пароля следует применять значение, сформированное с использованием всей информации, находящейся на компакт-диске. Полученное значение будет аналогом контрольной суммы или, что более правильно, хеш-значением (сжатым образом) всей информации компакт-диска, что исключить неправомерные действия по подмене содержимого компакт-диска или отдельных его частей. Хеш-функция обладает следующими достоинствами: хеш-функция чувствительна к всевозможным изменениям в информации M, таким как вставки, выбросы, перестановки и т.п.; задача подбора массива информации, который обладал бы

требуемым значением хеш-функции $h(M')$ равным $h(M)$ ($h(M') = h(M)$), вычислительно неразрешима; вероятность того, что значения хеш-функций двух различных массивов информации (вне зависимости от их длин) совпадут, ничтожно мала.

В качестве алгоритма получения хеш-функции можно применять ГОСТ Р34.11-94 или алгоритм безопасного хеширования SHA.

Применение интеллектуальных специализированных носителей позволяет упростить алгоритм идентификации, так как носитель сам обеспечивает безопасность записанной на нем информации и не позволяет считать или модифицировать хранящиеся на нем данные. Кроме этого следует отметить, что при работе с такими носителями необходимо применение оригинальных технических средств, специально разработанных для взаимодействия с подобными носителями информации. Информация на таких носителях также может храниться в преобразованном виде, что обеспечивает дополнительную защиту. Таким образом, возможно обеспечение безопасного легального управления объектом с разграничением форм доступа различных структур с внесением всех изменений в специальный реестр.

Совершенно другая ситуация складывается при управлении элементами объекта с использованием вспомогательных мобильных (подключаемых) АРМ. В данном случае первоначально необходимо идентифицировать подключаемую аппаратуру, а затем уже действовать по описанному выше алгоритму. Идентификация должна проходить в автоматическом режиме для того, чтобы исключить противоправные действия или непреднамеренные ошибки операторов. Наиболее безопасным и надежным алгоритмом идентификации двух устройств является алгоритм «запрос-ответ» [28-32].

В данной модели для инициирования работы необходимо передать в систему управления идентификационные данные подключаемого вспомогательного пульта управления. Система определяет параметры подключаемого устройства и формирует случайный запрос (для исключения возможности предопределенного ответа). Подключаемый пульт управления в соответствии с определенным алгоритмом преобразует полученное значение запроса и возвращает в систему управления. Исходя из результатов преобразований и полученной от подключаемого устройства информации, формируется разрешение или отказ в подключении вспомогательного пульта управления. При необходимости обеспечения безопасности передаваемой в каналах связи информации аппаратура системы управления может инициировать процедуру выработки сеансовых параметров преобразования информации для того, чтобы все данные, циркулирующие в линиях связи, были представлены в измененном (неявном) виде. Для формирования параметров сеанса связи и последующего преобразования передаваемой информации целесообразно применять гибридную схему, представляющую собой модификацию протокола распределения ключей Диффи-Хэллмана и алгоритма преобразования данных по ГОСТ 28147-89 в режиме гаммирования с обратной связью. Применение подобной схемы позволит сократить время на подготовительные операции и всего лишь незначительно уменьшить скорость передачи данных. Модель подобной системы приведена на рис. 1.

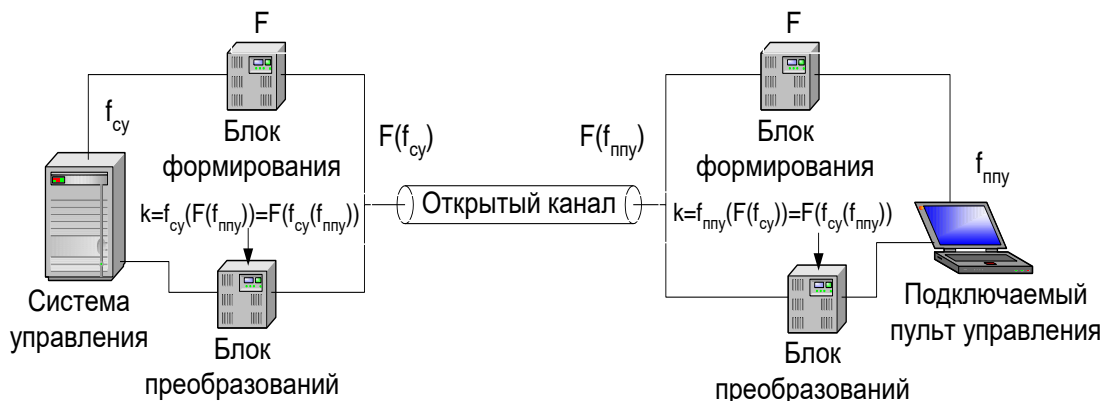


Рис. 1. Модель информационной системы

Согласно приведенному алгоритму обе стороны обладают некоторой уникальной информацией (система управления f_{cy} , подключаемый пульт управления $f_{ппу}$). Для формирования параметров сеанса связи уникальная информация преобразуется в блоках формирования при помощи специальной функции F ($F(f_{cy})$, $F(f_{ппу})$) и передается противоположной стороне (параметры функции F изменяются при каждом сеансе связи, при этом необходимо синхронизировать приемо-передающие устройства на обеих сторонах). Противоположная сторона с использованием своей уникальной информации вырабатывает требуемый параметр связи k и применяет его в блоке преобразования для осуществления безопасной передачи информации в течение всего сеанса связи. Преимуществом данной системы является отсутствие необходимости конфиденциальной передачи какой-либо информации по защищенному каналу связи.

Результаты исследований имеют важное прикладное значение в области создания безопасных систем дистанционного обучения на базе имеющихся и разрабатываемых протоколов безопасного обмена и распределения информации. Дальнейшие работы будут направлены на развитие общетеоретической концепции построения систем дистанционного обучения с целью построения формализованной математической модели процесса обеспечения и разграничения доступа, позволяющей применить существующий математический аппарат для описания происходящих образовательных процессов.

Библиографический список

1. Грибунин В.Г., Костюков В.Е., Мартынов А.П., Николаев Д.Б., Фомченко В.Н. Стеганографические системы. Цифровые водяные знаки: Учебно-методическое пособие // Под ред. В.Г. Грибунина. Саров: ФГУП "РФЯЦ-ВНИИЭФ", 2016. 295 с.
2. Бабанов Н.Ю., Мартынов А.П., Николаев Д.Б., Фомченко В.Н., Новиков А.В. Виртуальная интерактивная система формирования и отработки управляющей информации. Вестник НГИЭИ. 2016. № 4 (59). С. 15-29.
3. Мартынов А.П., Николаев Д.Б., Мартынова И.А. Криптографические системы и метод факториального сжатия информации // Известия института инженерной физики, 2016. № 4(42). С.54-57.
4. Грибунин В.Г., Костюков В.Е., Мартынов А.П., Николаев Д.Б., Фомченко В.Н. Современные методы обеспечения безопасности информации в атомной энергетике. Монография – Саров: ФГУП «РФЯЦ-ВНИИЭФ», 2014, 63бс.
5. Пат. 2623894, 07.09.2017. Способ преобразования данных с равновероятностной инициализацией / Мартынов А.П., Мартынова И.А., Марунин М.В., Николаев Д.Б., Фомченко В.Н. Патент на изобретение № 2623894 от 07.09.2017.
6. Пат. 2408924, 10.01.2011. Устройство обеспечения безопасного формирования параметров / Курочкин А.А., Мартынов А.П., Николаев Д.Б., Фомченко В.Н. Патент на изобретение № 2408924 от 10.01.2011.
7. Пат. 2256053. Устройство управления электронными замками / Гончаров С.Н., Мартынов А.П., Николаев Д.Б., Фомченко В.Н. Заявл. 20.01.2003, опубл. 10.07.2005. Бюл. № 19.
8. Мартынов А.П., Волков К.О., Николаев Д.Б. Обеспечение безопасного взаимодействия компонентов интегрированной системы. Инновации в условиях развития информационно-коммуникационных технологий. 2008. № 1. С. 136-138.
9. Борисенков И.А., Дороненков М.Н., Николаев Д.Б., Мартынов А.П., Фомченко В.Н. Концепция информационного взаимодействия участников процесса разработки наукоемкой продукции. В сборнике: Информатизация образования-2014 Материалы Международной научно-практической конференции. Редколлегия: Е.В. Данильчук, А.Н. Сергеев, Борисова Н.В., С.Н. Касьянов, Л.Ю. Кравченко, К.А. Попов, Е.М. Филиппова. 2014. С. 176-178.
10. Немченко И.А., Николаев Д.Б. Обеспечение безопасности качественной составляющей информации с использованием стохастических алгоритмов. Вестник Рязанского государственного университета имени С.А. Есенина. 2011. № 35. С. 126.

11. Мартынов А.П., Николаев Д.Б., Соколов С.Ю. Концептуальные основы построения систем обеспечения взаимной аутентификации объектов // Известия института инженерной физики, 2008. № 10(10). С.6-9.
12. Костюков В.Е., Николаев Д.Б., Мартынов А.П., Фомченко В.Н. Динамическая аутентификация группы разнородных объектов. Информатизация образования. 2014. Т. 2014. С. 260.
13. Сплюхин Д.В., Николаев Д.Б. Анализ новейших требований ФСТЭК и общие решения существующих проблем защиты информационных систем. Сборник материалов X-ой Всероссийской молодежной научно-инновационной школы «Математика и математическое моделирование». 2016. С. 28-29.
14. Грибунин В.Г., Костюков В.Е., Мартынов А.П., Николаев Д.Б., Фомченко В.Н. Стеганографические системы. Критерии и методическое обеспечение: Учебно-методическое пособие // Под ред. В.Г. Грибунина. Саров: ФГУП «РФЯЦ-ВНИИЭФ», 2016. 324 с.
15. Шишкин Г., Николаев Д. Селекторы цифровых команд. Часть 1 // Компоненты и технологии. 2009. № 6 (95). С.96-101.
16. Пат. 2154722. Электронный кодовый замок / Мартынов А.П., Николаев Д.Б., Шишкин Г.И. Заявл. 27.08.1999, опубл. 20.08.2000. Бюл. № 23.
17. Шишкин Г., Николаев Д. Селекторы цифровых команд. Часть 2 // Компоненты и технологии. 2009. № 8 (97). С.112-116.
18. Пат. 2402810, 27.01.2009. Устройство защиты от несанкционированного доступа к информации / Курочкин А.А., Мартынов А.П., Николаев Д.Б., Фомченко В.Н. Патент на изобретение № 2402810 от 27.01.2009.
19. Пат. 2497277, 27.10.2013. Способ сжатия двоичных данных в виде структурированных информационных блоков / Николаев Д.Б., Мартынов А.П. Патент на изобретение № 2497277 от 27.10.2013.
20. Пат. 2630429, 18.11.2016. Голихин М.В., Способ преобразования низкоэнтропийных сообщений / Мартынов А.П., Николаев Д.Б., Фомченко В.Н. Патент на изобретение № 2630429 от 18.11.2016.
21. Одинцов М.В., Сплюхин Д.В., Николаев Д.Б. Исследование вопросов оптимизации параметров защищенности информации. Сборник материалов IX-ой Всероссийской молодежной научно-инновационной школы «Математика и математическое моделирование». 2015. С. 75.
23. Гончаров С., Николаев Д., Никитин В., Писецкий В. Схемотехническая реализация автомата. Компоненты и технологии. 2013. № 2 (139). С. 126-128.
24. Шишкин Г., Николаев Д. Селекторы цифровых команд. Часть 3 // Компоненты и технологии. 2009. № 9 (98). С.116-120.
25. Пат. 2224075. Электронно-механический замок / Мартынов А.П., Николаев Д.Б., Сычев В.Т., Шишкин Г.И. Заявл. 28.11.2001, опубл. 20.02.2004. Бюл. № 5.
26. Шишкин Г., Николаев Д. Селекторы цифровых команд. Часть 4 // Компоненты и технологии. 2009. № 11 (100). С.102-106.
27. Васильев Р.А., Николаев Д.Б. Анализ возможностей применения голосовой идентификации в системах разграничения доступа к информации. Научный результат. Информационные технологии. 2016. Т. 1. № 1. С. 48-57.
28. Астайкин А.И., Мартынов А.П., Николаев Д.Б., Фомченко В.Н. Методы и средства обеспечения программно-аппаратной защиты информации. Научно-техническое издание / Саров, 2015.
29. Пат. 2546238, 07.10.2013. Способ транслитерационного преобразования информации и передачи ее по каналам связи / Грибунин В.Г., Мартынов А.П., Николаев Д.Б., Фомченко В.Н. Патент на изобретение № 2546238 от 07.10.2013.

30. Пат. 2554525, 07.10.2013. Способ преобразования информации с синхронной сменой инициализирующих последовательностей в блоках, соединенных каналом связи с неопределенным периодом смены / Фомченко В.Н., Мартынов А.П., Николаев Д.Б., Масыгин А.М., Мартынов А.А. Патент на изобретение № 2554525 от 07.10.2013.
31. Пат. 2416169, 20.07.2009. Способ аутентификации объекта / Мартынов А.П., Николаев Д.Б. Патент на изобретение № 2416169 от 20.07.2009.
32. Пат. 2506633, 26.07.2012. Устройство хранения данных (варианты) / Фомченко В.Н., Николаев Д.Б., Гончаров С.Н., Жердин О.А., Силаев А.Г., Марунин М.В., Моксяков А.Н. Патент на изобретение № 2506633 от 26.07.2012.
33. Грибунин В.Г., Костюков В.Е., Мартынов А.П., Николаев Д.Б., Фомченко В.Н. Стеганографические системы. Атаки, пропускная способность каналов и оценка стойкости: Учебно-методическое пособие // Под ред. В.Г. Грибунина. Саров: ФГУП "РФЯЦ-ВНИИЭФ", 2015.

A.A. Evstifeev K.D. Ermakov

**SOFTWARE COMPLEX FOR AUTOMATION OF THE REMOTE TESTING
PROCESS**

Sarov Institute of Physics and Technology - branch of NRNU MEPhI

Variants of delimiting access to the elements of the information resource and ensuring the conflict-free operation of all involved elements during the automation of remote testing processes are considered.

Keywords: request-response, identity, password, transformation, access control, unique features.

ПОДХОД К ПОСТРОЕНИЮ БЕЗОПАСНОГО АЛГОРИТМА ДОСТУПА К ИНФОРМАЦИИ ПРИ ПРОВЕДЕНИИ ПРЕДМЕТНЫХ ОЛИМПИАД

Саровский физико-технический институт – филиал НИЯУ МИФИ

Приведены результаты разработки основных элементов алгоритмов распределения и защиты информации при проведении предметных олимпиад школьников.

Ключевые слова: криптоанализ, криптографические методы, олимпиада, пароль, преобразование, разграничение доступа, уникальные особенности.

Предметные олимпиады школьников проводятся в несколько этапов. Это школьные, городские (районные), областные (республиканские, краевые), зональные и Всероссийские олимпиады. Олимпиады проводятся практически по всем основным предметам. Подготовку заданий для них осуществляют специалисты очень высокой квалификации, которые представляют, как правило, более высшую инстанцию и пользуются всеобщим уважением и доверием. Между тем, никто не застрахован от того, что в обслуживающем персонале не будет людей, которые стремятся использовать полученную информацию для своей выгоды или для получения ряда преимуществ определенным кругом участников олимпиад. При данных условиях утечка информации может произойти на конечных этапах сортировки, в процессе доведения заданий или в процессе проведения олимпиад (когда они проводятся в разных местах с определенным временным сдвигом). В этом случае защиту информации целесообразно осуществлять по определенным алгоритмам методами шифрования с использованием криптографических функций и протоколов.

Целью проведения исследований является разработка основных элементов алгоритмов распределения и защиты информации при проведении предметных олимпиад школьников. Ограничения на основных элементах вводятся потому, что сложность и доказуемость правильности применения ряда современных криптографических алгоритмов и протоколов в целом выходит далеко не только за рамки школьной программы, но и за рамки программ большинства ВУЗов страны и в данном докладе их оценить трудно [1-4].

Анализ возможных алгоритмов подготовки олимпиадных заданий и проведения олимпиад показывает, что утечка информации может произойти на трех основных этапах:

- 1) на этапе подготовки и формирования олимпиадных заданий;
- 2) на этапе доставки заданий к пунктам назначения;
- 3) на этапе проведения олимпиад (когда олимпиады проводятся в нескольких пунктах в разное время, которое изменяется от нескольких часов до нескольких суток).

Для достижения поставленной цели необходимо решить две взаимосвязанные задачи:

- 1) на основе сложившейся практики проведения олимпиад необходимо предложить новый модифицированный алгоритм их проведения, включающий все этапы от подготовки заданий до подведения итогов;
- 2) разработать для данного алгоритма механизмы распределения и защиты информации.

Начало исследований криптографических методов и алгоритмов защиты информации было положено в работе [5]). В работе рассмотрены:

- простейшие методы шифрования Скитала, Атбаш, шифр Цезаря и шифрующие диски;
- два метода криптоанализа: частотный анализ и метод вероятных слов.

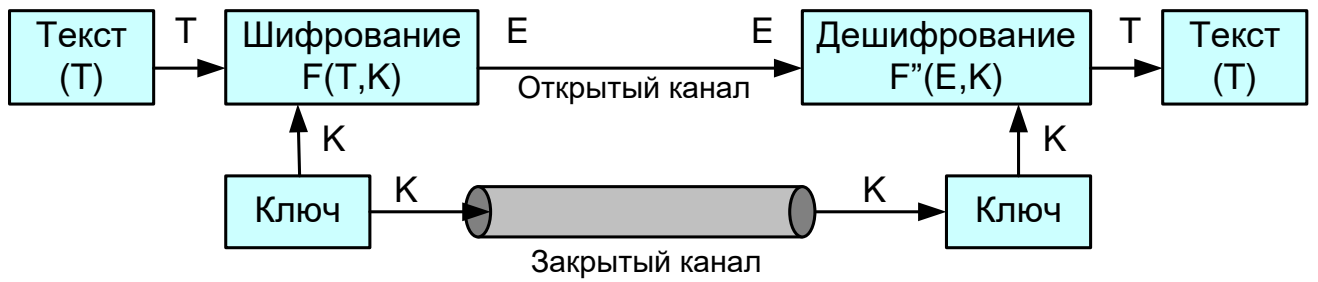


Рис. 1. Схема процесса шифрования и расшифрования

По полученным статистическим данным были подсчитаны вероятности встречаемости букв в текстах на русском и английском языках.

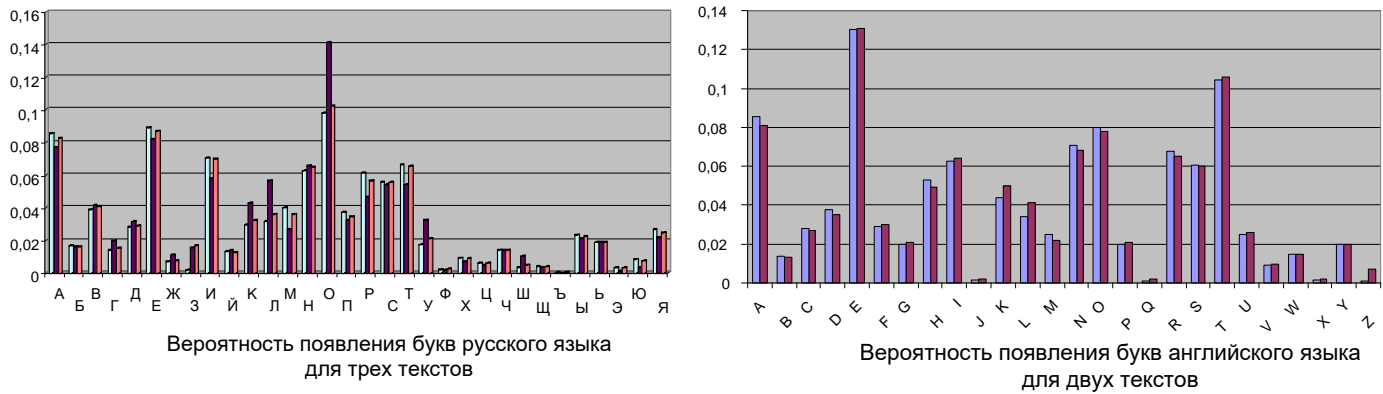


Рис. 2. Вероятности появления букв в текстах

Выполнен криптоанализ одного из текстов:

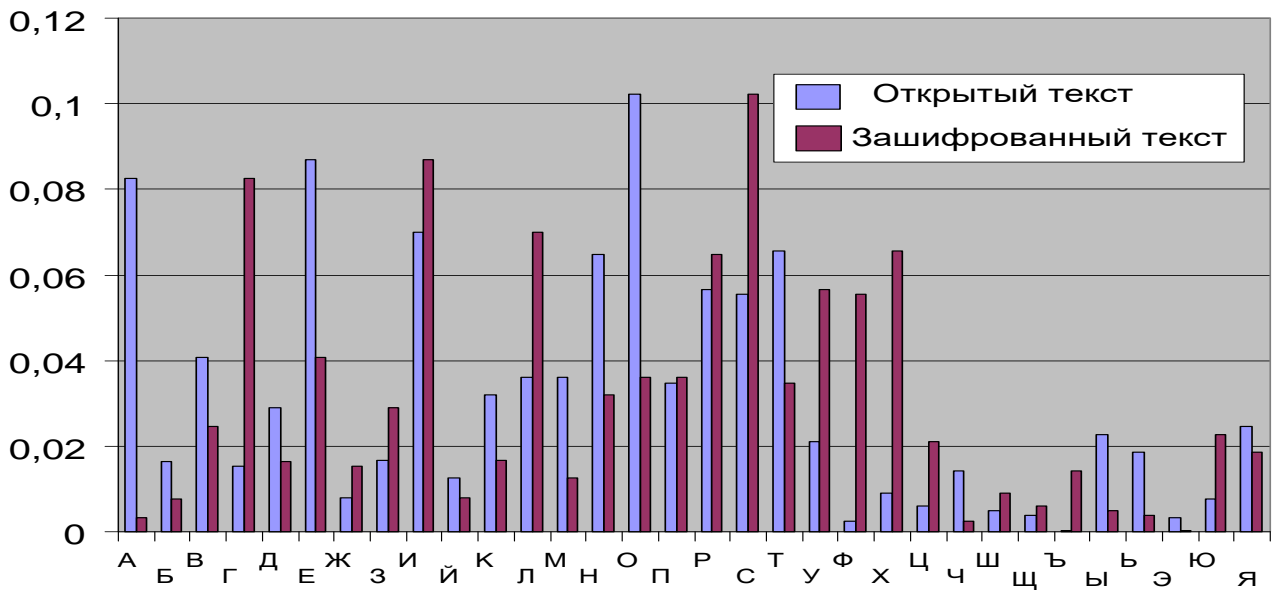


Рис. 3. Результат анализа

В нем было показано, что простейшие криптографические методы не обеспечивают защиту серьезных результатов, к которым относятся и олимпиадные задания. В работе рассмотрены методы, основанные на частотном и вероятностном криптоанализе информации. Основные выводы были подтверждены на Нижегородской сессии молодых ученых [6], где спектр криптографических методов шифрования и их криптоанализа был расширен до известных методов, широко использовавшихся в середине прошлого столетия. Основное программное обеспечение, написано на языке Паскаль.

При проведении дальнейших исследований рассмотрены вопросы обеспечения конфиденциальности информации с применением преобразующих алгоритмов [7], применения современных протоколов для обеспечения целостности информации [8] и разработки алгоритмов обмена конфиденциальными данными [9,10].

В результате сделаны следующие выводы:

- 1) все шифры, рассмотренные в работах [1, 2] нельзя применять для наших задач;
- 2) новейшие протоколы и алгоритмы обмена информацией [5, 6] будут слишком сложными для практической реализации, дорогими и избыточными для нашего применения;
- 3) для решения нашей задачи идеально подходит отечественный Алгоритм криптографического преобразования по ГОСТ 28147-89:
 - для шифрования информации – в режиме гаммирования с обратной связью;
 - для обеспечения целостности – в режиме выработки имитовставки (контрольной суммы);
- 4) все каналы связи лучше исключить и передавать информацию на CD дисках. Ключи расшифровки можно в нужный момент передавать по телефону.

Это исключит необходимость реализации закрытого канала связи и упростит криптографический протокол.

Опираясь на полученные результаты в докладе подробно рассмотрен процесс подготовки и проведения предметных олимпиад от начала выработки заданий до их непосредственного проведения.

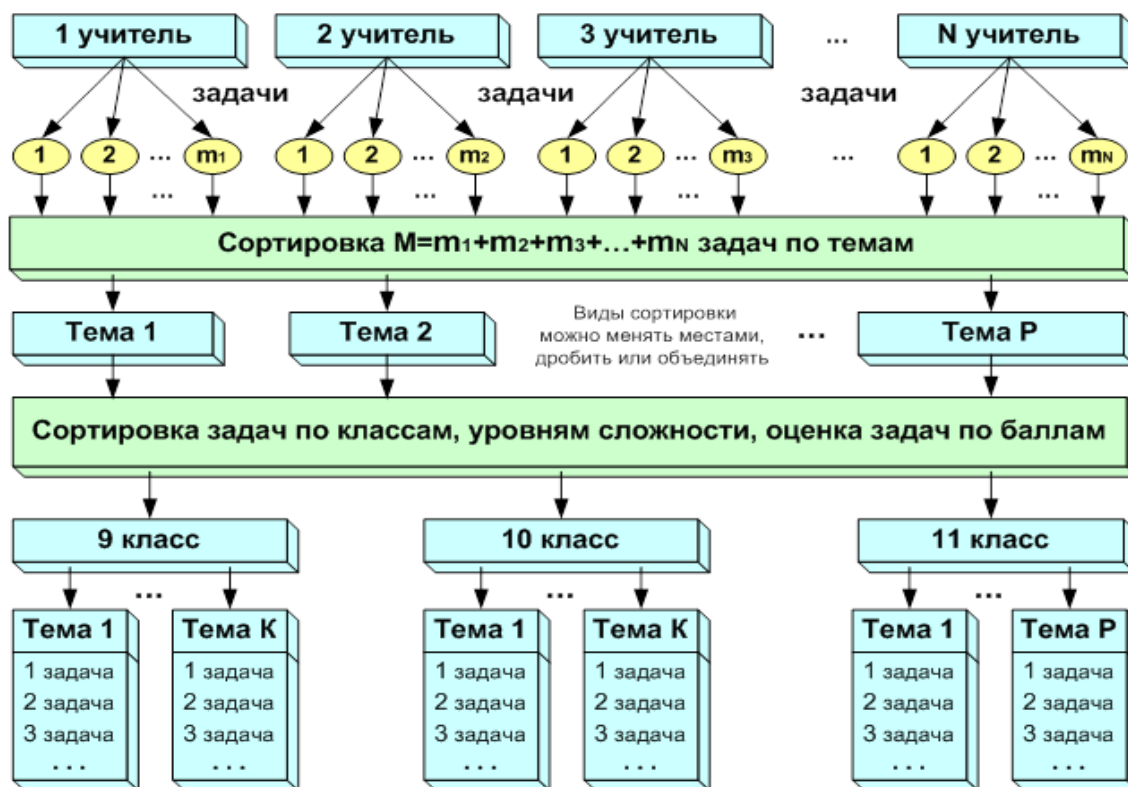


Рис. 4. Сортировка задач по тематикам

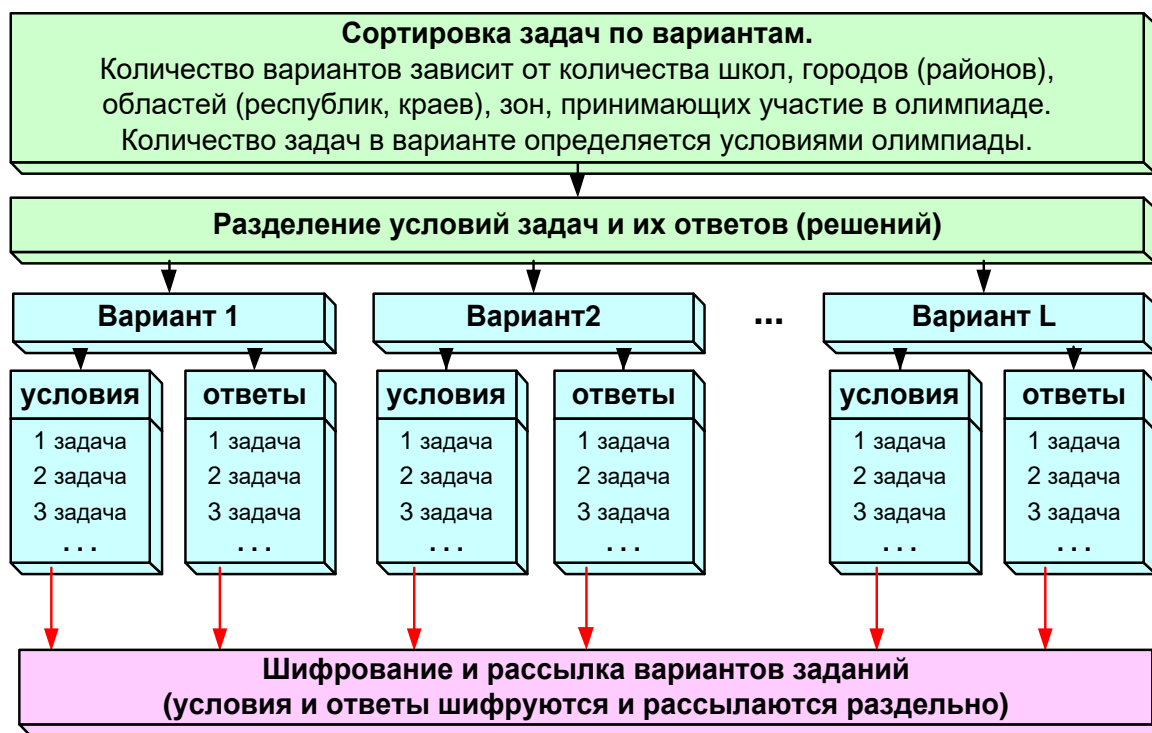


Рис. 5. Процесс сортировки задач по вариантам

Предложена алгоритмическая основа для формализации и автоматизации основных этапов подготовки, классификации, сортировки, доведения и непосредственного использования олимпиадных заданий. Причем доведение заданий предложено осуществлять в обезличенном и зашифрованном виде:

1. Шифрование текстов задач и ответов осуществляется по ГОСТ 28147-89 в режиме гаммирования с обратной связью;
2. Если зашифрованную информацию передавать на CD дисках, то ключи расшифрования можно передавать в последний момент по телефону или в конверте. Это исключит необходимость закрытого канала связи;

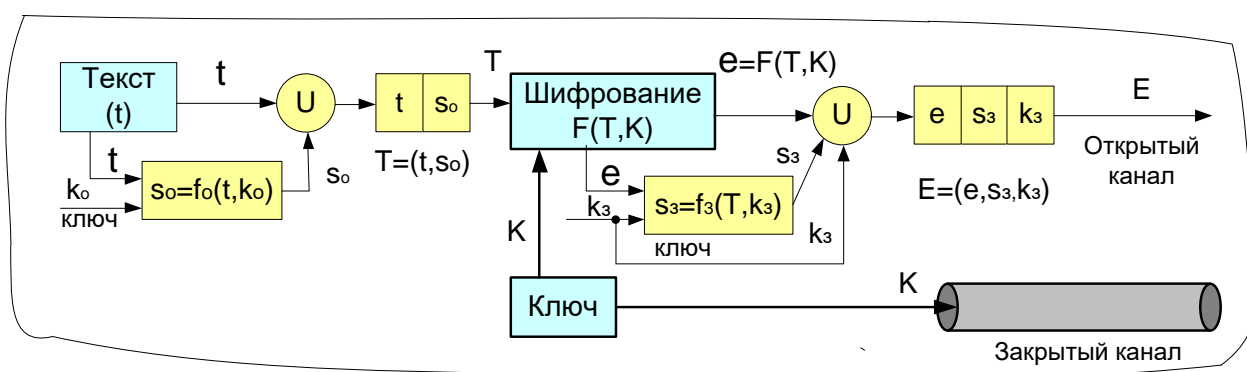


Рис. 6. Алгоритм шифрования

3. Для защиты от изменения используется режим выработки имитовставки (криптографической контрольной суммы) по ГОСТ 28 147-89;
4. Кроме шифрования и расшифрования вариантов заданий необходимо обеспечить их целостность (защиту от изменения). Защиту от изменения лучше обеспечить как для зашифрованных, так и для незашифрованных вариантов заданий.

Расшифровка заданий может осуществляться непосредственно перед началом олимпиады с учетом их размножения. Расшифровка правильных ответов может осуществляться непосредственно перед их проверкой.

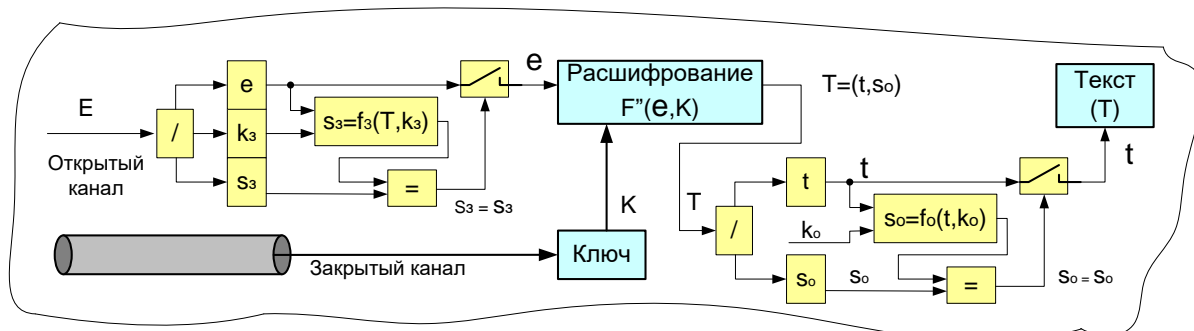


Рис. 7. Алгоритм расшифрования

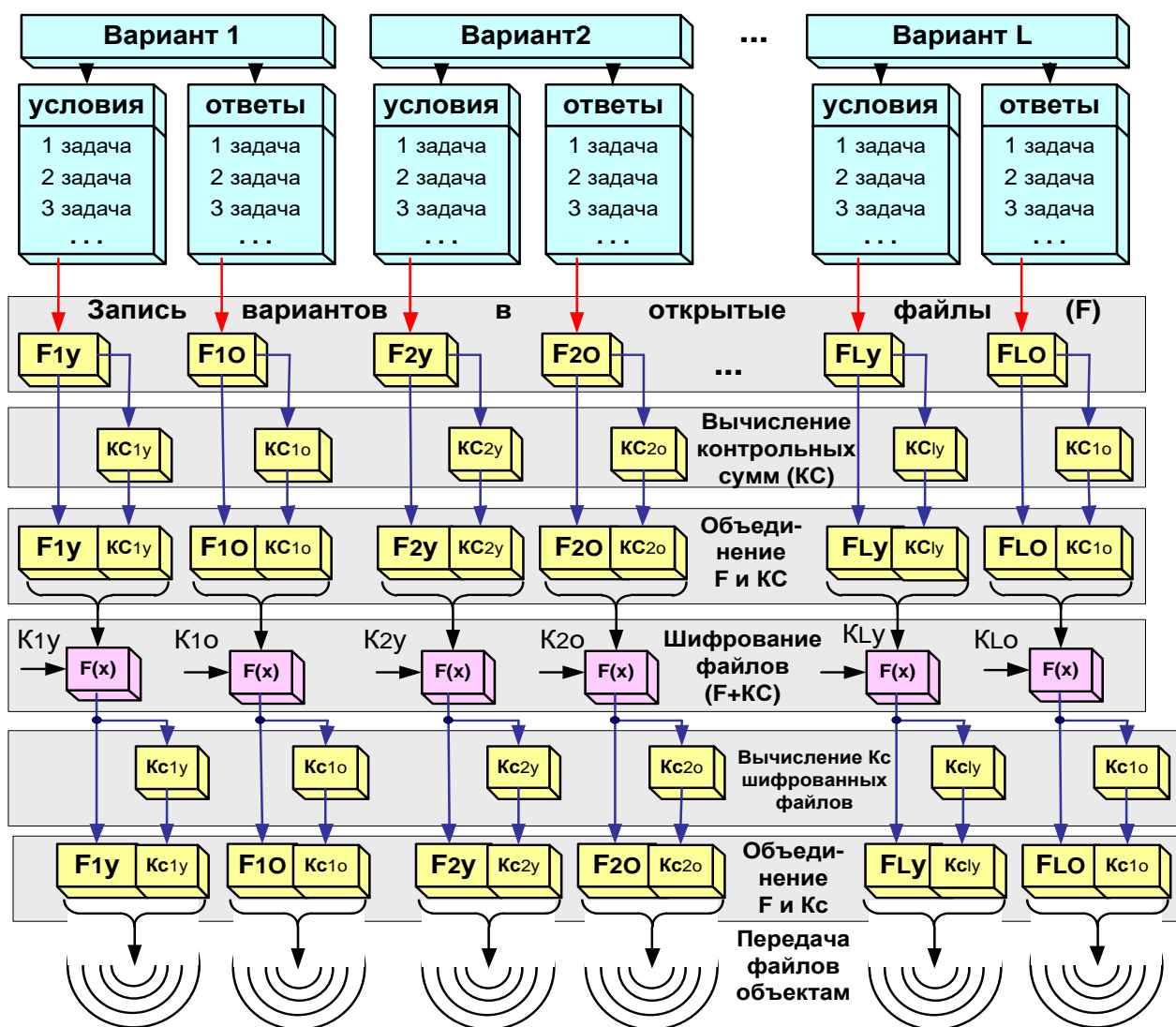


Рис. 8. Процесс зашифрования и отправки файлов с вариантами

Временные задержки в доставке ключей расшифрования для файлов с заданиями и ответами определяются вышестоящим органом, который осуществлял подготовку и рассылку заданий.

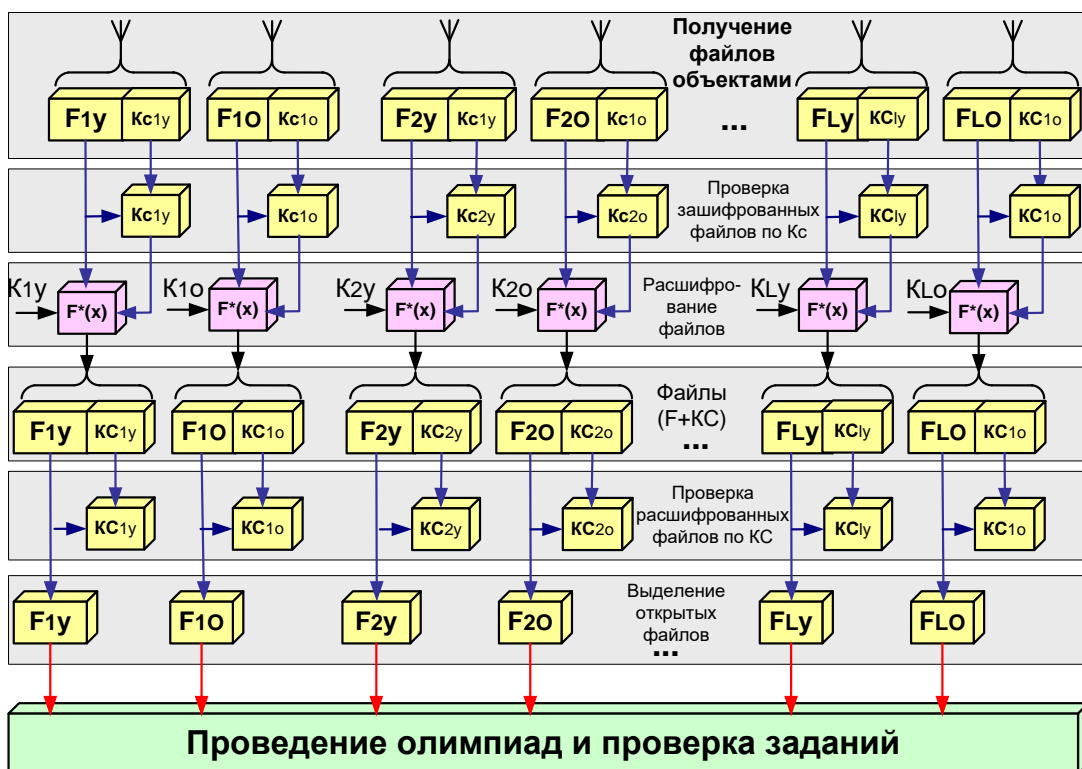


Рис. 9. Процесс получения и расшифрования файлов с вариантами

Научная значимость работы заключается в том, что на основе сложившейся практики проведения олимпиад предложен новый защищенный модифицированный алгоритм их организации и проведения, включающий этапы от подготовки заданий до подведения итогов;

Практическая значимость работы заключается в том, что для данного алгоритма предложены механизмы распределения и защиты, обеспечивающие секретность и целостность информации.

Особенностью предложений является то, что рассматриваемые алгоритмы могут позволить сокрытие конкретного наполнение вариантов задачами даже от самого вышестоящего органа, позволяют обеспечить проверку правильности и целостности вариантов заданий на всех основных этапах, как в процессе подготовки, так и в процессе проведения олимпиад. Кроме этого, рассматриваемые подходы к обеспечению безопасности могут быть распространены и на электронные устройства доступа [11-13].

Библиографический список

1. Алферов А.П., Зубов А.Ю., Кузьмин А.С., Черемушкин А.В. Основы криптографии. Учебное пособие. – М.: Гелиос АРВ, 2001. – 480 с., ил.
2. Мартынов А.П., Фомченко В.Н. Основы криптографии. Под редакцией А.И. Астайкина. Саров: ФГУП «РФЯЦ-ВНИИЭФ». 2006, 452 с.
3. Девид Кан. Взломщики кодов. Перевод с английского А. Ключевского. – «Секретная папка». ЗАО Изд-во Центрполиграф. 2000 – 473с.
4. Шнайер Б. Прикладная криптография. Протоколы, алгоритмы, исходные тексты на языке СИ. – М.: Издательство ТРИУМФ, 2003 – 816 с.: ил.
5. Мартынова И. Методы защиты результатов физических экспериментов. Тезисы докладов международной конференции VII Харитоновские чтения, стр.50-51, 2007.
6. Мартынова И.А., Николаев Д.Б. Исследование вопросов обеспечения безопасности современных алгоритмов преобразования информации, 12 Нижегородская сессия молодых ученых и студентов (математические науки), 2007.

7. Мартынова И.А., Николаев Д.Б. Обеспечение конфиденциальности информации с применением преобразующих алгоритмов. XXVI межрегиональная научно-техническая конференция «Проблемы эффективности и безопасности функционирования сложных технических и информационных систем» (г. Серпухов Московской области, 28 – 29 июня 2007 года).
8. Гончаров С.Н., Кузин Д.С., Мартынова И.А. Применение современных протоколов для обеспечения целостности информации. I Всероссийская научно-практическая конференция «Информационные технологии в образовании, науке и производстве» (г. Серпухов Московской области, 2 – 6 июля 2007 года).
9. Кузин Д.С., Мартынова И.А., Николаев Д.Б. Алгоритм обмена конфиденциальными данными. VI научно-техническая конференция «Молодежь в науке» (г. Саров, 30, 31 октября – 1 ноября 2007 года).
10. Кузин Д.С., Мартынова И.А., Николаев Д.Б. Алгоритм обмена конфиденциальными данными. Школа «Математика и математическое моделирование», СарФТИ, Саров, ноябрь 2007 года.
11. Пат. 2256053. Устройство управления электронными замками / Гончаров С.Н., Мартынов А.П., Николаев Д.Б., Фомченко В.Н. Заявл. 20.01.2003, опубл. 10.07.2005. Бюл. № 19.
12. Пат. 2224075. Электронно-механический замок / Мартынов А.П., Николаев Д.Б., Сычев В.Т., Шишкин Г.И. Заявл. 28.11.2001, опубл. 20.02.2004. Бюл. № 5.
13. Пат. 2154722. Электронный кодовый замок / Мартынов А.П., Николаев Д.Б., Шишкин Г.И. Заявл. 27.08.1999, опубл. 20.08.2000. Бюл. № 23.

I.A. Martynova, D.B. Nikolaev

**APPROACH TO BUILDING A SAFE ALGORITHM OF ACCESS TO
INFORMATION WHEN CONDUCTING SUBJECT OLYMPIADS**

Sarov Institute of Physics and Technology - branch of NRNU MEPhI

The results of the development of the main elements of algorithms for the distribution and protection of information during subject Olympiads for schoolchildren are presented.

Keywords: cryptanalysis, cryptographic methods, Olympiad, password, transformation, access control, unique features.

ИССЛЕДОВАНИЕ СВОЙСТВ ДВУХКЛЮЧЕВОГО АЛГОРИТМА, ИСПОЛЬЗУЕМОГО В СИСТЕМАХ ДИСТАНЦИОННОГО ОБРАЗОВАНИЯ

Саровский физико-технический институт – филиал НИЯУ МИФИ

Представлены результаты исследования свойств двухпараметрических схем, которые обеспечивают формирование конфиденциальных параметров и как следствие являются основой протоколов и систем идентификации в системах дистанционного образования.

Ключевые слова: адаптивное воздействие, вычислительная модель, двухпараметрические системы, идентификация, злоумышленник, полиномиальные алгоритмы.

Дистанционное образование является одной из наиболее перспективных ветвей развития дополнительного обучения, особенно при введении ограничений на посещение образовательных учреждений. Дистанционное обучение подразумевает наличие высококачественной информационной коммуникации между участниками учебного процесса при динамическом изменении состава и участников и характеристик оборудования, участвующего во взаимодействии. При этом должна достигаться целостность информационной системы и конфиденциальность циркулирующей в ней информации. Такой комплексный подход оптимальным образом реализуется в криптографических системах с многопараметрическими характеристиками, когда конфиденциальные ключи системы взаимосвязаны с открытыми ключами. Рассмотрим наиболее известные функции и криптосистемы, предложенные в криптографии с открытым ключом.

Диффи и Хеллман предложили исключительно простой способ использования дискретных алгоритмов для обмена секретными ключами между пользователями компьютерной или телефонной сети с применением только открытых сообщений. Назовем его протоколом Диффи-Хеллмана. Полагаем, что всем пользователям известны A и p . Пользователь i случайным образом выбирает целое число X_i из диапазона ($1 \leq X_i \leq p$) и держит его в секрете. Далее он вычисляет

$$Y_i = A^{X_i} \pmod{p} \quad (1)$$

Пользователь не держит Y_i в секрете, а помещает его в «справочник» доступный для всех.

В дальнейшем, если пользователи i и j захотят установить секретную связь, пользователь i возьмет из «справочника» Y_j и с помощью своего секретного X_i вычислит

$$Z_{ij} = (Y_j)^{X_i} = (A^{X_j})^{X_i} = A^{X_i \cdot X_j} \pmod{p} \quad (2)$$

Таким же образом и пользователь j вычисляет Z_{ji} . Однако $Z_{ij} = Z_{ji}$ и пользователи i и j могут с этого момента использовать Z_{ij} как секретный ключ в классической криптосистеме.

Если криптоаналитик смог бы решить задачу вычисления дискретных логарифмов, то он мог бы, получив из справочника Y_i и Y_j решить уравнение

$$X = \log_a Y_i \quad (3)$$

и вычислить Z_{ij} подобно пользователю i . По всей видимости он не может определить Z_{ij} другим способом. Схема для обмена секретными ключами между пользователями с использованием задачи вычисления дискретных алгоритмов приведена на рис. 1.

В этом способе в качестве возможной базовой функции Диффи и Хеллманом предложена функция дискретного возведения в степень:

$$F(X) = A^x \pmod{p}, \quad (4)$$

где X – целое число от 1 до $(p-1)$ включительно, p – очень большое простое число, A – целое число ($1 \leq A \leq p$) степени которого равняются (в некотором порядке) 1, 2, ..., $(p-1)$. В алгебре такое A называется примитивным элементом конечного поля $GF(p)$, и известно, что такие A всегда существуют.

Если $Y = A^x$, то

$$X = \log_a(Y), \quad (5)$$

а задача обращения называется задачей нахождения дискретных логарифмов. Даже при очень больших $p = 2^{1000}$, можно легко вычислить $F(X)$ путем возведения в квадрат и умножения.

Для $p = 2^{1000}$ вычисление $F(X)$ для любого целого X из диапазона ($1 \leq X \leq p$) потребуется менее 2000 операций умножения по mod p .

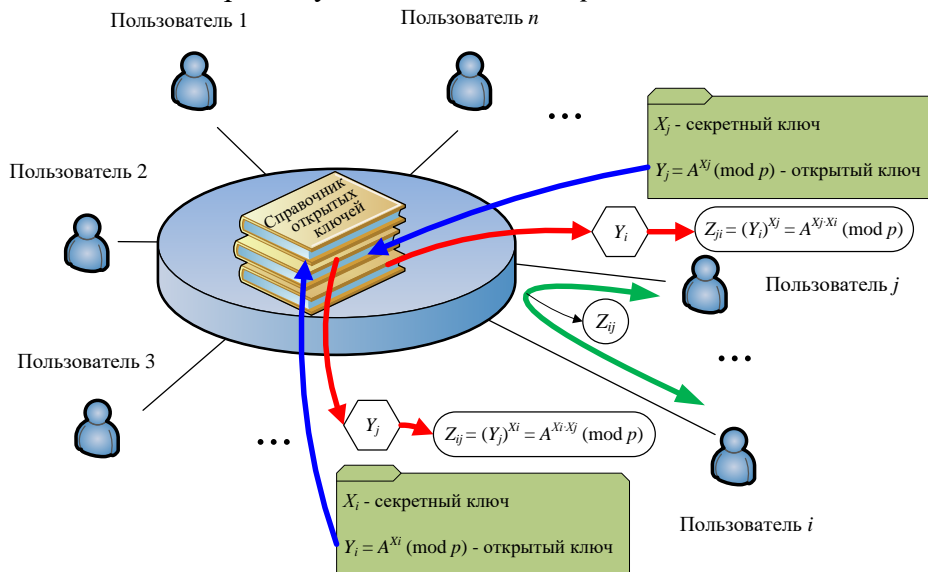


Рис. 1. Схема для обмена секретными ключами между пользователями с использованием задачи вычисления дискретных алгоритмов

Дискретные логарифмы сложно вычисляются при условии, когда не только p велико, но и $(p-1)$ имеет большой простой множитель (лучше всего, если это другое простое число, умноженное на 2). При этом дополнительном условии лучшие известные алгоритмы для нахождения дискретных логарифмов требуют примерно корень квадратный из p умножений по mod p (более точное соотношение приведено ниже) по сравнению с примерно $2 \cdot \log_2(p)$ умножений при дискретном возведении в степень.

Если p выбрано правильно, то извлечение логарифмов по mod p потребует проведения предварительных вычислений, по объему пропорциональных

$$L(p) \approx e^{\sqrt{\ln(p) \ln \ln(p)}} \quad (6)$$

после осуществления которых отдельные логарифмы могут быть вычислены достаточно просто. С помощью функции $L(p)$ оценивается также время, требуемое для разложения на множители составного числа аналогичной величины. При указанных условиях функция дискретного возведения в степень действительно является односторонней.

Односторонняя функция RSA с потайным ходом предложена Ривестом, Шамиром и Адлманом в 1978 году (по первым буквам RSA). Односторонняя функция RSA является одной из самых надежных в криптографии с открытым ключом. Она необычайно проста, но для ее описания необходимы некоторые сведения из теории чисел.

Пусть НОД (i, n) – наибольший общий делитель целых чисел i и n , которые одновременно не равны нулю, например, НОД (12,18) = 6. Для каждого положительного целого n функция Эйлера $\Phi(n)$ определяется как число положительных целых i , не

превосходящих n и таких, что $\text{НОД}(i, n) = 1$ (при $n=1$, по определению, $\Phi(1) = 1$), например, $\Phi(6)=2$, т.к. из всех $1 \leq i < 6$ лишь $i=1$ и $i=5$ дают $\text{НОД}(i, n) = 1$.

Очевидно, что для простого числа p имеем $\Phi(p) = (p-1)$, а для двух неравных простых чисел p и q

$$\Phi(pq) = (p-1)(q-1). \quad (7)$$

Теорема Эйлера гласит, что для любых целых чисел x и n ($x < n$)

$$X^{\Phi(n)} = 1 \pmod{n}. \quad (8)$$

При условии, что $\text{НОД}(x, n) = 1$. Например $5^2 = 1 \pmod{6}$.

Согласно Эвклиду, если e и m удовлетворяют условиям $0 < e < m$ и $\text{НОД}(m, e) = 1$, то существует единственное d , такое что $0 < d < m$ и

$$d \cdot e = 1 \pmod{m}. \quad (9)$$

Алгоритмом ассимметричного шифрования RSA является просто дискретное возведение в степень:

$$Fz(X) = X^e \pmod{n}, \quad (10)$$

где X – положительное целое, не превосходящее $n = pq$, потайной ход $z = \{p, q, e\}$, p и q – большие неравные числа, такие, что $\Phi(n) = (p-1)(q-1)$ имеет большой простой множитель, а e – положительное целое, не превосходящее $\Phi(n)$, для которого $\text{НОД}(e, \Phi(n)) = 1$.

Алгоритм Ez для быстрого вычисления Fz находится легко: это просто метод возведения в квадрат и умножения. Публикация этого алгоритма для практического использования сводится к публикации значений n и e .

Обратная функция имеет вид:

$$Fz^{-1}(Y) = Y^d \pmod{n}, \quad (11)$$

где d – единственное положительное целое меньше n и удовлетворяющее условию

$$d \cdot e = 1 \pmod{\Phi(n)}. \quad (12)$$

Алгоритм Dz для вычисления обратной функции тоже находится легко – это тоже метод возведения в квадрат и умножения. Показатель d , необходимый при расшифровании, находится с помощью алгоритма Эвклида, вычисляющего $\text{НОД}(e, \Phi(n))$.

Тот факт, что функция (11) действительно является функцией, обратной к (12), показывается следующим образом. Равенство $de = 1 \pmod{\Phi(n)}$ эквивалентно (в обычной целочисленной арифметике)

$$d \cdot e = \Phi(n)Q + 1 \quad (13)$$

для некоторого Q . Из формул (10) и (13) получаем $e = (\Phi(n) + 1) / d$

$$(X^e)^d = X^{\Phi(n)Q+1} \pmod{n} = X^{\Phi(n)} \cdot X^Q \pmod{n} = X \pmod{n}, \quad (14)$$

где в последнем равенстве использована теорема Эйлера (8). В ней требуется выполнение условия $\text{НОД}(x, n) = 1$. Однако, в частном случае, когда n является произведением двух неравных простых чисел, равенство (8) верно для всех положительных целых X , меньших n .

Равенство (14) показывает, что операция возведения числа в степень d (по модулю n) обратна по отношению к операции возведения в степень e (по модулю n). Стойкость криптосистемы RSA основана на допущении о том, что любой способ обращения функции Fz эквивалентен разложению $n=pq$ на множители, т.е. зная способ обращения Fz можно лишь с небольшими дополнительными затратами разложить на множители.

Лучшие известные алгоритмы решения задачи дискретных логарифмов (по $\text{mod } p$) и лучшие алгоритмы разложения n на множители требуют при p примерно равном n приблизительно одинакового количества вычислений. Поэтому функция RSA и функция Диффи-Хеллмана имеют на сегодняшний день примерно равные основания считаться односторонними.

Для анализа криптосистемы RSA можно использовать алгоритм Шэнкса. Алгоритм Шэнкса позволяет получить решение показательного уравнения $g^x = a$ в некоторой группе G или найти порядок циклической группы, если известен диапазон возможных значений.

Поиском эффективных способов разложения целых чисел на множители занимаются уже очень давно. Эта задача интересовала выдающихся ученых в области теории чисел. Вероятно, Ферма был первый, кто предложил представить разлагаемое число N в виде разности квадратов $N = x^2 - y^2$, а затем, вычисляя $(N, x - y)$, попытаться найти нетривиальный делитель N . Он же предложил и способ, позволяющий найти требуемое представление. Если разлагаемое число имеет два не очень отличающиеся по величине множителя, этот способ позволяет определить их быстрее, чем простой перебор делителей. Лежандр обратил внимание на то, что при таком подходе достаточно получить сравнение

$$x^2 \equiv y^2 \pmod{N}. \quad (15)$$

Напомним, что каждому иррациональному числу ξ может быть поставлена в соответствие бесконечная последовательность целых чисел $[b_0, b_1, b_2, \dots]$, называемая его непрерывной дробью. Это сопоставление строится следующим образом

$$x_0 = \xi, b_i = [x_i], x_{i+1} = \frac{1}{x_i - b_i}, i = 0, 1, 2, \dots$$

Лежандр доказал, что непрерывная дробь квадратичной иррациональности периодична. Если раскладывать в непрерывную дробь число $\xi = \sqrt{N}$, то возникающие в процессе разложения числа x_i , имеют вид $x_i = \frac{\sqrt{N} + P_i}{Q_i}$ с целыми P_i, Q_i причем всегда $0 < P_i < \sqrt{N}, 0 < Q_i < 2\sqrt{N}$. С каждой непрерывной дробью можно связать последовательность рациональных чисел, так называемых подходящих дробей, $\frac{A_i}{B_i}, i \geq 0$, вычисляемых по правилам $A_{i+1} = b_{i+1}A_i + A_{i-1}, B_{i+1} = b_{i+1}B_i + B_{i-1}, i \geq 0, A_0 = b_0, B_0 = A_{-1} = 1, B_{-1} = 0$ и стремящихся к разлагаемому числу. Если в непрерывную дробь разлагается число $\xi = \sqrt{N}$, то справедливо соотношение

$$A_{i-1}^2 \equiv (-1)^i Q_i \pmod{N}. \quad (16)$$

В 1971 г. Шенкс предложил использовать сравнения (16) для конструирования чисел, удовлетворяющих (15). Если вычисления проводить до тех пор, пока при четном i не получится $Q_i = R^2$ при некотором целом R , то пара чисел $\langle A_{i-1}, R \rangle$ будет удовлетворять (15) и с ее помощью можно надеяться получить разложение N на простые множители.

Предлагаемый авторами способ, базирующийся на алгоритме Шенкса, предназначен для решения показательного уравнения $g^x = a$ в некоторой группе G или для нахождения порядка циклической группы, если известен диапазон возможных значений. В частности, данный способ может применяться для нахождения числа точек алгебраической кривой или для разложения составного числа (впрочем, не очень длинного), а также для нахождения ключа шифра, если ключи образуют группу или полугруппу.

Способ решения показательного уравнения $g^x = a$ на основе алгоритма Шэнкса.

Вход. Группа G , образующая g , ее образ a , максимально возможное значение для показателя x .

Выход. Показатель x .

Способ.

1. Найти целое число $r \approx \sqrt{x_m}$.

2. Найти элемент g^r .

3. Построить базу данных из пар $\{(i, g^{ir})\}$ («большой шаг») для $1 < i < r$ и отсортировать ее по компоненту g^{ir} .

4. Вычислять ag^j («маленький шаг») для $j=0, 1, 2, \dots$ до совпадения с каким-либо элементом базы данных. (При совпадении получается равенство $g^{ir+j}=a$, число попыток не более $r+1$.)
5. Результат: $x = ir + j$.

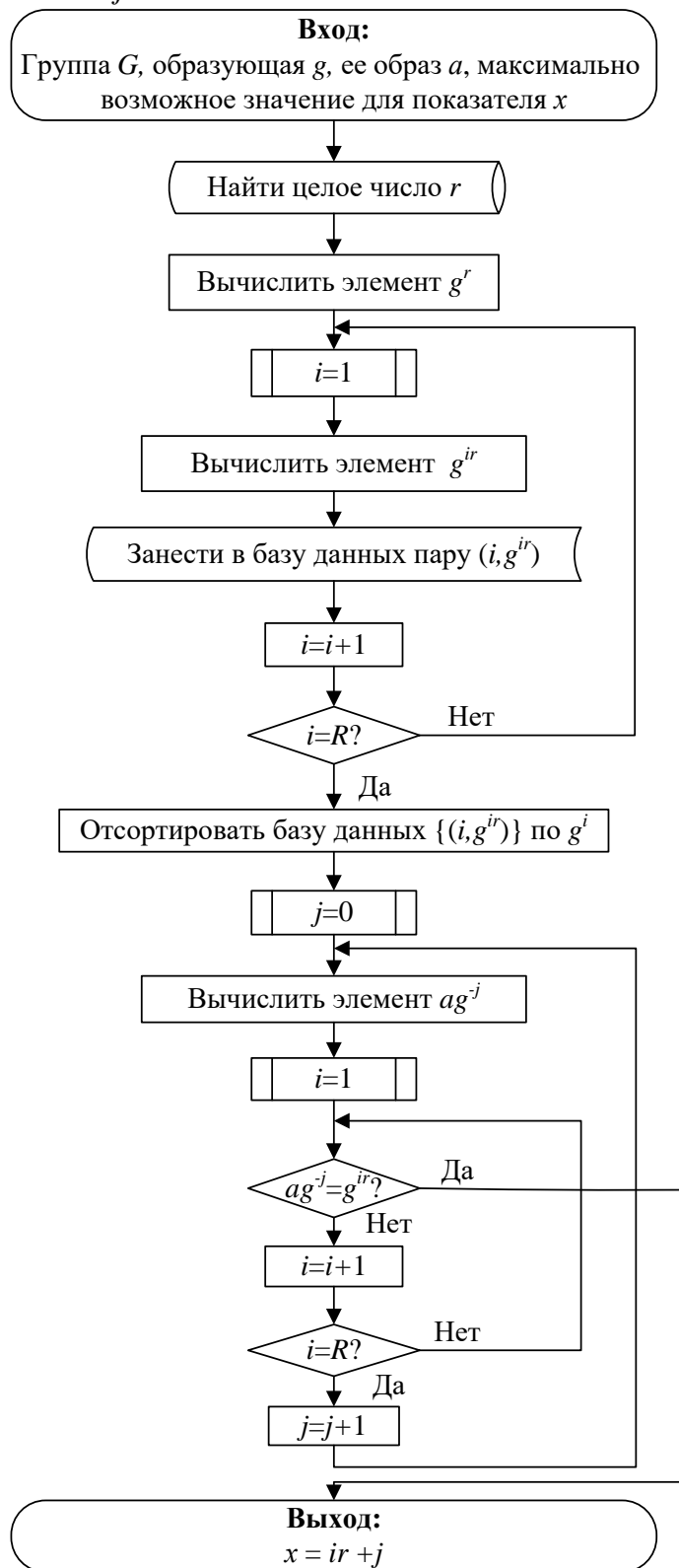


Рис. 2. Блок-схема предложенного способа решения уравнения $g^x = a$ в группе G

Алгоритм требует $r = O(\sqrt{x_m})$ умножений в группе G при построении базы данных, $O(r \log r) = O(\sqrt{x_m} \log x_m)$ операций при сортировке базы данных и не более $r = O(\sqrt{x_m})$

умножений в группе G на шаге 4. Если сложность умножения в группе равна $O(1)$, то итоговая сложность алгоритма равна $O(\sqrt{x_m} \log x_m)$ требуемая память имеет такую же оценку.

Очевидно, что задача разложения числа n усложняется с ростом n . Если n имеет несколько простых делителей, то достаточно разложить число n на два сомножителя (не обязательно простых) и затем решать более простые задачи разложения для найденных коротких сомножителей. Количество разложений числа n на два сомножителя растет с увеличением числа простых делителей, поэтому задача разложения наиболее сложна, когда n является произведением двух различных простых чисел.

Задача разложения такого числа n полиномиально эквивалентна вычислению значения функции Эйлера $\varphi(n)$, так как по известному разложению легко находится $\varphi(n)$, а по известному значению $\varphi(n)$ делители числа n находятся как корни квадратного уравнения $x^2 - (n - \varphi(n) + 1)x + n = 0$. Следовательно, задача разложения полиномиально эквивалентна задаче логарифмирования в следующей формулировке: для числа g , взаимно простого с n , найти показатель $\varphi(n)$ такой, что $a^{\varphi(n)} = 1 \pmod{n}$. Здесь $1 < n - \varphi(n) < 2\sqrt{n}$, и разложение предложенным способом имеет сложность $O(\sqrt[4]{n} \log n)$.

Предложенный способ на основе алгоритма Шэнкса может быть использован и при нахождении числа точек N эллиптической кривой над конечным полем из q элементов. По теореме Хассе $|q + 1 - N| \leq 2\sqrt{q}$. В этом случае решается уравнение $NP = P_\infty$, алгоритм имеет сложность $O(\sqrt[4]{q} \log q)$.

Проведенные исследования показали, что предложенный способ имеет более низкую сложность выполнения вычисления по сравнению с существующими аналогами. Сравнительные качественные характеристики сложности существующих и предложенного способов приведены на рис. 3.

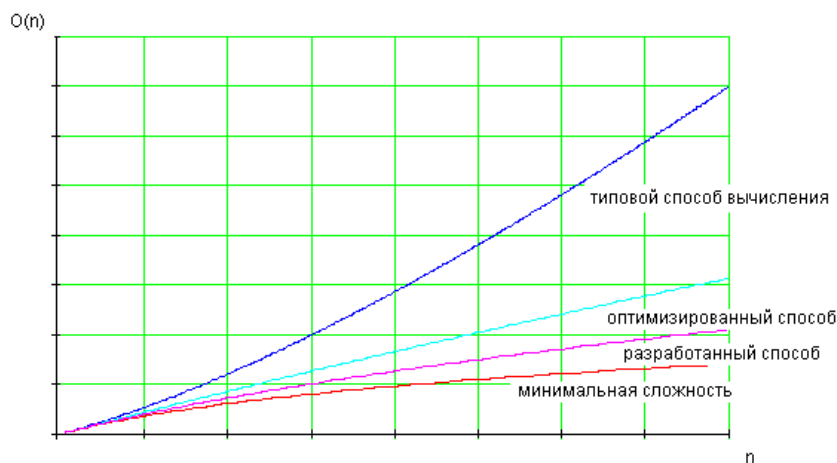


Рис. 3. Сравнительные качественные характеристики сложности существующих и предложенного способов

Библиографический список

1. Борисенков И.А., Дороненков М.Н., Николаев Д.Б., Мартынов А.П., Фомченко В.Н. Концепция информационного взаимодействия участников процесса разработки наукоемкой продукции. В сборнике: Информатизация образования-2014. 2014. С. 176-178.
2. Немченко И.А., Николаев Д.Б. Обеспечение безопасности качественной составляющей информации с использованием стохастических алгоритмов. Вестник Рязанского государственного университета имени С.А. Есенина. 2011. № 35. С. 126.
3. Мартынов А.П. и др. Концептуальные основы построения систем обеспечения взаимной аутентификации объектов // Известия института инженерной физики, 2008. № 10(10). С.6-9 .
4. Костюков В.Е., Николаев Д.Б., Мартынов А.П., и др. Динамическая аутентификация группы разнородных объектов. Информатизация образования. 2014. Т. 2014. С. 260.
5. Сплюхин Д.В., Николаев Д.Б. Анализ новейших требований ФСТЭК и общие решения существующих проблем защиты информационных систем. Сборник материалов X-ой Всероссийской молодежной научно-инновационной школы «Математика и математическое моделирование». 2016. С. 28-29.
6. Грибунин В.Г., Костюков В.Е., Мартынов А.П., Николаев Д.Б., Фомченко В.Н. Стеганографические системы. Критерии и методическое обеспечение: Учебно-методическое пособие // Под ред. В.Г. Грибунина. Саров: ФГУП «РФЯЦ-ВНИИЭФ», 2016, 324 с.
7. Шишкин Г., Николаев Д. Селекторы цифровых команд. Часть 1 // Компоненты и технологии. 2009. № 6 (95). С.96-101.
8. Пат. 2154722. Электронный кодовый замок / Мартынов А.П., Николаев Д.Б., Шишкин Г.И. Заявл. 27.08.1999, опубл. 20.08.2000. Бюл. № 23.
9. Шишкин Г., Николаев Д. Селекторы цифровых команд. Часть 2 // Компоненты и технологии. 2009. № 8 (97). С.112-116.
10. Пат. 2402810, 27.01.2009. Устройство защиты от несанкционированного доступа к информации / Курочкин А.А., Мартынов А.П., Николаев Д.Б., Фомченко В.Н. Патент на изобретение № 2402810 от 27.01.2009.
11. Грибунин В.Г., Костюков В.Е., Мартынов А.П., Николаев Д.Б., Фомченко В.Н. Стеганографические системы. Цифровые водяные знаки: Учебно-методическое пособие // Под ред. В.Г. Грибунина. Саров: ФГУП «РФЯЦ-ВНИИЭФ», 2016, 295 с.
12. Бабанов Н.Ю., Мартынов А.П. и др. Виртуальная интерактивная система формирования и обработки управляющей информации. Вестник НГИЭИ. 2016. № 4 (59). С. 15-29.
13. Мартынов А.П., Николаев Д.Б., Мартынова И.А. Криптографические системы и метод факториального сжатия информации // Известия института инженерной физики, 2016. № 4(42). С.54-57.
14. Грибунин В.Г., Костюков В.Е., Мартынов А.П., Николаев Д.Б., Фомченко В.Н. Современные методы обеспечения безопасности информации в атомной энергетике. Монография – Саров: ФГУП «РФЯЦ-ВНИИЭФ», 2014, 636с.
15. Пат. 2623894, 07.09.2017. Способ преобразования данных с равновероятностной инициализацией / Мартынов А.П., Мартынова И.А., Марунин М.В., Николаев Д.Б., Фомченко В.Н. Патент на изобретение № 2623894 от 07.09.2017.
16. Пат. 2256053. Устройство управления электронными замками / Гончаров С.Н., Мартынов А.П., Николаев Д.Б., Фомченко В.Н. Заявл. 20.01.2003, опубл. 10.07.2005. Бюл. № 19.
17. Пат. 2224075. Электронно-механический замок / Мартынов А.П., Николаев Д.Б., Сычев В.Т., Шишкин Г.И. Заявл. 28.11.2001, опубл. 20.02.2004. Бюл. № 5.
18. Грибунин В.Г., Костюков В.Е., Мартынов А.П., и др. Стеганографические системы. Атаки, пропускная способность каналов и оценка стойкости: Учебно-методическое пособие // Под ред. В.Г. Грибунина. Саров: ФГУП "РФЯЦ-ВНИИЭФ", 2015.

Y.D. Zharkova, K.D. Ermakov

**RESEARCH OF THE PROPERTIES OF THE TWO-KEY ALGORITHM USED
IN DISTANCE EDUCATION SYSTEMS**

Sarov Institute of Physics and Technology - branch of NRNU MEPhI

The paper presents the results of a study of the properties of two-parameter schemes, which provide the formation of confidential parameters and, as a result, are the basis of protocols and identification systems in distance education systems.

Keywords: adaptive action, computational model, two-parameter systems, identification, intruder, polynomial algorithms.

ПРИМЕНЕНИЕ АСИММЕТРИЧНЫХ АЛГОРИТМОВ ДЛЯ ОБЕСПЕЧЕНИЯ ИНФОРМАЦИОННОЙ БЕЗОПАСНОСТИ

Саровский физико-технический институт – филиал НИЯУ МИФИ

Дано описание применения асимметричных алгоритмов для обеспечения информационной безопасности личности в условиях рисков и вызовов современного общества.

Ключевые слова: алгоритм, данные, генерация, модификация, открытый параметр, преобразование, угроза.

Информационная безопасность личности обеспечивается применением механизмов защиты, связанных с методами контроля и идентификации. Одним из наиболее часто используемых подходов к контролю и идентификации является асимметричный алгоритм преобразования данных RSA. Этот алгоритм может использоваться для реализации безопасности личных данных, электронной цифровой подписи, распределения сессионных параметров и во многих других случаях. В условиях возрастающего количества рисков и вызовов информационное общество должно соответствовать современным требованиям по безопасности. Поскольку циркуляция личных данных осуществляется в каждом сегменте информационного общества, то и внешним злоумышленным воздействиям личные данные подвергаются также очень часто. Следовательно, оценка эффективности применяемых средств обеспечения безопасности является актуальной задачей. Вместе с тем, оценка позволяет сделать вывод о возможности применения децентрализованного алгоритма с целью использования в нем более коротких параметров при той же криптографической стойкости, либо улучшения его криптографических характеристик является актуальной задачей [1].

В основу алгоритма с открытым параметром RSA положена задача умножения и разложения простых чисел на множители, которая является вычислительно односторонней задачей [2].

Предположим, сторона B хочет послать стороне A сообщение M .

Сообщением являются целые числа лежащие от 0 до $n-1$, т.е $M \in D = Z_n$.

Алгоритм:

- взять открытый параметр (e, n) стороны A ;
- взять исходный текст M ;
- передать сообщение $P_A(M) = M^e$.

(1)

Алгоритм:

- принять сообщение C ;
- применить конфиденциальный параметр (d, n) для восстановления сообщения

$$S_A(C) = C^d. \quad (2)$$

Уравнения (1) и (2), на которых основана схема RSA, определяют взаимно обратные преобразования множества Z_n .

Доказательство.

Действительно, $\forall M \in Z_n. P(S(M)) = S(P(M)) = M^{ed} \pmod n$.

Докажем, что $M^{ed} \equiv M \pmod p \forall M$.

Возможны два случая.

1) $M \neq 0 \pmod p$ и, поскольку числа e и d являются взаимно обратными относительно умножения по модулю $\varphi(n) = (p-1)(q-1)$, т.е. $ed = 1 + k(p-1)(q-1)$ для

некоторого целого k , имеем $M^{ed} = M(M^{p-1})^{k(q-1)} \pmod{p} = M(1)^{k(q-1)} \pmod{p} = M \pmod{p}$, где второе тождество следует из теоремы Ферма.

2) $M \equiv 0 \pmod{p}$, тогда $M^{ed} \equiv M \pmod{p}$.

Таким образом, при всех M выполняется равенство $M^{ed} \equiv M \pmod{p}$.

Аналогично можно показать, что $M^{ed} \equiv M \pmod{q} \forall M$.

Таким образом, из Китайской теоремы об остатках $M^{ed} \equiv M \pmod{n} \forall M$.

Поскольку генерация параметров происходит значительно реже операций, реализующих преобразования, а также создание и проверку цифровой подписи, задача вычисления $a = b^e \pmod{n}$ представляет основную вычислительную сложность. Эта задача может быть разрешена с помощью алгоритма быстрого возведения в степень. Таким образом, для вычисления $M^e \pmod{n}$ требуется $O(\ln e)$ операций умножения по модулю.

Доказательство:

- представим e в виде: $e = e_k \cdot 2^k + e_{k-1} \cdot 2^{k-1} + \dots + e_1 \cdot 2 + e_0$, где $e_k = 1, e_i \in \{0,1\}$;
- положим $M_0 = 1$;
- вычислим $M_i = M_{i-1}^2 \cdot M^{e_{k-i}}$ для $i=1, \dots, k$;
- найденное M_k и будет искомым значением $M^e \pmod{n}$.

Так как каждое вычисление на шаге 2 требует не более трёх умножений по модулю n и этот шаг выполняется $k \leq \log_2 e$ раз, то сложность может быть оценена величиной $O(\ln e)$.

Чтобы проанализировать время выполнения операций с открытым и конфиденциальным параметрами, предположим, что открытый параметр (e, n) и конфиденциальный параметр (d, n) удовлетворяют соотношениям $\log_2 e = O(1), \log_2 d \leq \beta$. Тогда в процессах их применения выполняются соответственно $O(1)$ и $O(\beta)$ умножений по модулю.

Таким образом, время выполнения операций растёт с увеличением количества ненулевых битов в двоичном представлении открытой экспоненты e . Чтобы увеличить скорость преобразования, значение e часто выбирают равным 17, 257 или 65537 – простым числам, двоичное представление которых содержит лишь две единицы: $17 = 0x11, 257 = 0x101, 65537 = 0x10001$ (простые числа Ферма).

По эвристическим оценкам длина конфиденциальной экспоненты d , нетривиальным образом зависящей от открытой экспоненты e и модуля n , с большой вероятностью близка к длине n . Поэтому восстановление данных идёт медленнее, чем преобразование, а проверка подписи быстрее, чем сама подпись.

Размер параметра в алгоритме RSA связан с размером модуля n . Два числа p и q , произведением которых является модуль, должны иметь приблизительно одинаковую длину, поскольку в этом случае найти сомножители (факторы) сложнее, чем в случае, когда длина чисел значительно различается. Например, если предполагается использовать 768-битный модуль, то каждое число должно иметь длину приблизительно 384 бита. Если два числа чрезвычайно близки друг к другу или их разность близка к некоторому предопределенному значению, то возникает потенциальная угроза безопасности, однако такая вероятность – близость двух случайно выбранных чисел – незначительна.

1. Возьмем $M = (p+q)/2$

2. При $p < q$, имеем $0 \leq M - n^{1/2} \leq (q-p)^{2/8p}$.

Поскольку $p = M (\pm (m^2 - n)^{1/2})$, то значения p и q можно легко найти, если разность $p - q$ достаточно мала.

Оптимальный размер модуля определяется требованиями безопасности: модуль большего размера обеспечивает большую безопасность, но и замедляет работу алгоритма RSA. Длина модуля выбирается, в первую очередь, на основе значимости защищаемых данных и необходимой стойкости защищенных данных и, во вторую очередь, – на основе оценки возможных угроз.

В математической основе предложенных модификаций лежит алгоритм преобразования RSA по 32 битному параметру. В дополнение к преобразованию блоков данных путем возведения в степень используются дополнительные меры для улучшения характеристик текста, а также для ускорения преобразования файлов [3-5].

Модификация №1 работает следующим образом (рисунок 1). Из открытого файла в бинарном режиме считываются 1024 блока по 16 бит (либо меньше, в случае, если конец файла достигнут). Далее происходит перемешивание блоков с помощью побитовой операции «Исключающее или» (обозначим ее символом \wedge): $V1=V1\wedge V0$, $V2=V2\wedge V1\dots V1023=V1023\wedge V1022$, $V0=V0\wedge V1023$. Затем каждый блок преобразуется по алгоритму RSA (в результате каждый блок из 16 бит преобразуется в блок данных из 32 бит, т.к. основание равно 32 битам), и результат записывается в новый файл. Восстановление полученного файла происходит аналогичным образом, но в обратном порядке: из файла читаются 1024 блока данных по 4 бита, каждый блок восстанавливается конфиденциальным параметром RSA, в результате чего каждый блок преобразуется в двухбитный. Затем блоки перемешиваются в обратном порядке [6-10]. В итоге получаем исходный файл.

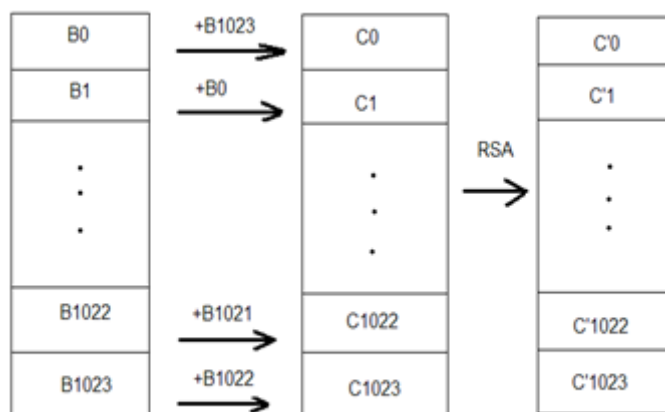


Рис. 1. Схема работы первой модификации

Данный алгоритм работает дольше, чем RSA, но, вследствие перемешивания блоков данных, возможно, текст имеет более приемлемые характеристики, чем при преобразовании обычным RSA [11-13].

Модификация №2 работает следующим образом (рис. 2). Из файла в двоичном режиме считываются блоки данных по 16 бит. Далее из считанного блока данных формируется основа параметра длиной 8 бит (берутся первые и последние 4 бита). Полный параметр формируется из основы ее дублированием (т.е. если основа параметра $k1$, то полный параметр представляет собой число $k1k1$, или $k1*256+k1$). Далее идет преобразование считанного блока данных с помощью операции «Исключающего ИЛИ» (если прочитанный блок X , $k1k1$ – параметр, то блок S формируется – $S = X\wedge k1k1$). Но помимо записи в файл текста идет добавочная запись преобразованного по алгоритму RSA параметра $k1$. И таким образом преобразуются каждые считанные 16 бит данных.

Восстановление происходит в обратном порядке – считывается блок данных в 16 бит и параметр. Параметр декодируется по алгоритму RSA, и по полученному параметру находится исходный блок данных ($S\wedge k1k1 = (X\wedge k1k1)\wedge k1k1 = X$). Операция повторяется для каждого считанного блока данных.

Алгоритм работает быстрее RSA, т.к. все операции занимают мало вычислительного времени, а при преобразовании параметра в степень открытой (конфиденциальной) экспоненты возводится число не в 16 бит длиной, а только 8.

На основе реализованной программы было произведено сравнение времени преобразования файла большого объема (порядка 700МБ) тремя алгоритмами (RSA и его модификациям). Результаты можно видеть в таблице.

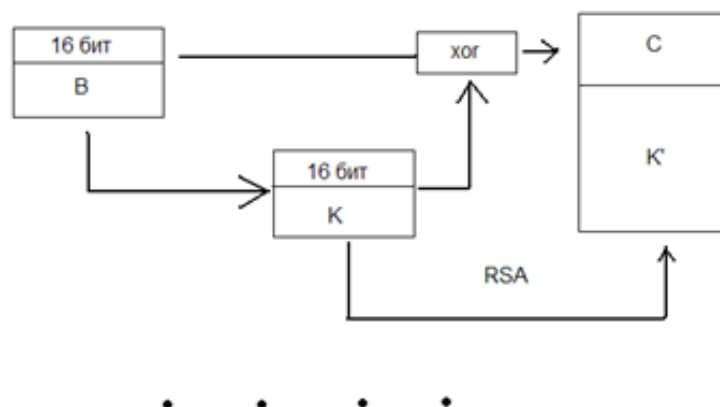


Рис. 2. Схема работы второй модификации

Как видно из таблицы, вторая модификация работает быстрее, чем алгоритм RSA, в то время как первая работает медленнее. Но первый алгоритм обеспечивает перемешивание блоков данных и, возможно, улучшает статистические характеристики текста [7]. Проведенный анализ показал целесообразность дальнейшего исследования предложенных алгоритмов. В случае неудовлетворительных результатов возможно изменение предложенных алгоритмов [14-18].

Таким образом, были разработаны две модификации алгоритма RSA с параметрическими последовательностями в 32 бита, которые реализованы в программном обеспечении, преобразующем файлы системы. Разработанные методы сравнивались по временной трудоемкости (результаты представлены в таблице). По результатам сравнения, вторая модификация работает быстрее алгоритма RSA [8-12].

Таблица

Время преобразования файла объемом 700 МБ	
Алгоритм	Время преобразования
RSA	215 секунд
Первая модификация	270 секунд
Вторая модификация	165 секунд

Оценка эффективности асимметричных алгоритмов для обеспечения безопасности личности в условиях рисков и вызовов современного общества позволила сделать следующие выводы: алгоритм RSA подвержен множеству внешних воздействий, и для обеспечения конфиденциальности передаваемой информации необходимо использовать параметрическую информацию большого объема; не во всех приложениях (аппаратных или программных) возможно использование параметров порядка 1024 бит и более, для решения подобных задач были разработаны модификации алгоритма RSA; при сравнении по временной трудоемкости разработанные алгоритмы сопоставимы с RSA (вторая модификация работает быстрее), но для более глубокого анализа необходимо осуществить дальнейшие исследования их криптографических характеристик.

Библиографический список

1. Бабанов Н.Ю., Мартынов А.П. и др. Виртуальная интерактивная система формирования и отработки управляющей информации. Вестник НГИЭИ. 2016. № 4 (59). С. 15-29.

2. Васильев Р.А., Николаев Д.Б. Анализ возможностей применения голосовой идентификации в системах разграничения доступа к информации. Научный результат. Информационные технологии. 2016. Т. 1. № 1. С. 48-57.
3. Немченко И.А., Николаев Д.Б. Обеспечение безопасности качественной составляющей информации с использованием стохастических алгоритмов. Вестник Рязанского государственного университета имени С.А. Есенина. 2011. № 35. С. 126.
4. Грибунин В.Г., Костюков В.Е., Мартынов А.П., Николаев Д.Б., Фомченко В.Н. Стеганографические системы. Цифровые водяные знаки: Учебно-методическое пособие // Под ред. В.Г. Грибунина. Саров: ФГУП "РФЯЦ-ВНИИЭФ", 2016. 295 с.
5. Борисенков И.А., Дороненков М.Н., Николаев Д.Б., Мартынов А.П., Фомченко В.Н. Концепция информационного взаимодействия участников процесса разработки наукоемкой продукции. В сборнике: Информатизация образования-2014 Материалы Международной научно-практической конференции. Редколлегия: Е.В. Данильчук, А.Н. Сергеев, Борисова Н.В., С.Н. Касьянов, Л.Ю. Кравченко, К.А. Попов, Е.М. Филиппова. 2014. С. 176-178.
6. Пат. 2546238, 07.10.2013. Способ транслитерационного преобразования информации и передачи ее по каналам связи / Грибунин В.Г., Мартынов А.П., Николаев Д.Б., Фомченко В.Н. Патент на изобретение № 2546238 от 07.10.2013.
7. Пат. 2623894, 07.09.2017. Способ преобразования данных с равновероятностной инициализацией / Мартынов А.П., Мартынова И.А., Марунин М.В., Николаев Д.Б., Фомченко В.Н. Патент на изобретение № 2623894 от 07.09.2017.
8. Пат. 2554525, 07.10. 2013. Способ преобразования информации с синхронной сменой инициализирующих последовательностей в блоках, соединенных каналом связи с неопределенным периодом смены / Фомченко В.Н., Мартынов А.П., Николаев Д.Б., Мясягин А.М., Мартынов А.А. Патент на изобретение № 2554525 от 07.10.2013.
9. Пат. 2402810, 27.01.2009. Устройство защиты от несанкционированного доступа к информации / Курочкин А.А., Мартынов А.П., Николаев Д.Б., Фомченко В.Н. Патент на изобретение № 2402810 от 27.01.2009.
10. Пат. 2416169, 20.07.2009. Способ аутентификации объекта / Мартынов А.П., Николаев Д.Б. Патент на изобретение № 2416169 от 20.07.2009.
11. Пат. 2506633, 26.07.2012. Устройство хранения данных (варианты) / Фомченко В.Н., Николаев Д.Б., и др. Патент на изобретение № 2506633 от 26.07.2012.
12. Пат. 2630429, 18.11.2016. Способ преобразования низкоэнтропийных сообщений / Мартынов А.П., Николаев Д.Б. и др. Патент на изобретение № 2630429 от 18.11.2016.
13. Пат. 2187612, 20.08.2002. Электронный кодовый замок / Мартынов А.П., Николаев Д.Б., Шишкин Г.И. Патент на изобретение № 2187612 от 08.05.2001.
14. Пат. 2408924, 10.01.2011. Устройство обеспечения безопасного формирования параметров / Курочкин А.А., Мартынов А.П., Николаев Д.Б., Фомченко В.Н. Патент на изобретение № 2408924 от 10.01.2011.
15. Пат. 2154722. Электронный кодовый замок / Мартынов А.П., Николаев Д.Б., Шишкин Г.И. Заявл. 27.08.1999, опубл. 20.08.2000. Бюл. № 23.
16. Пат. 2224075. Электронно-механический замок / Мартынов А.П., Николаев Д.Б., Сычев В.Т., Шишкин Г.И. Заявл. 28.11.2001, опубл. 20.02.2004. Бюл. № 5.
17. Пат. 2256053. Устройство управления электронными замками / Гончаров С.Н., Мартынов А.П. и др. Заявл. 20.01.2003, опубл. 10.07.2005. Бюл. № 19.
18. Одинцов М.В., Сплюхин Д.В., Николаев Д.Б. Исследование вопросов оптимизации параметров защищенности информации. Сборник материалов IX-ой Всероссийской молодежной научно-инновационной школы «Математика и математическое моделирование». 2015. С. 75.

A.A. Kazakov, A.M. Tarasov, I.A.Nikolaeva

**APPLICATION OF ASYMMETRIC ALGORITHMS TO ENSURE INFORMATION
SECURITY OF AN INDIVIDUAL**

Sarov Institute of Physics and Technology - branch of NRNU MEPhI

The article describes the use of asymmetric algorithms to ensure information security of a person in the face of risks and challenges of modern society.

Keywords: algorithm, data, generation, modification, open parameter, transformation, threat.

**ПРИМЕНЕНИЕ МЕТОДА МОНТГОМЕРИ ДЛЯ ИНЖЕНЕРНЫХ ПРИЛОЖЕНИЙ
ФИЗИКО-МАТЕМАТИЧЕСКИХ НАУК***Саровский физико-технический институт – филиал НИЯУ МИФИ*

Рассмотрен вопрос разработки быстрых алгоритмов выполнения модульных операций (умножения, сложения, возведение в степень) с использованием арифметики Монтгомери для выполнения указанных модульных операций в полях больших целых чисел в криптографических системах.

Ключевые слова: быстродействие, криптографические системы, модульные операции, программная реализация, сложность.

Защита информации различного назначения криптографическими средствами является бурно развивающейся областью прикладной науки. В настоящее время, в связи с бурным развитием информационных технологий, криптографические методы защиты информации применяются не только для защиты конфиденциальной информации государственной важности, но и для обеспечения информационной безопасности организаций и частных лиц.

Наибольшее распространение в настоящее время получают асимметричные криптографические системы, криптостойкость которых основана на трудоемкости вычисления обратных функций.

Асимметричные криптографические системы (рисунок 1) основаны на двухпараметрических системах, для прямого и обратного преобразования в которых используются разные параметры. Применение данного преобразования имеет неоспоримое преимущество в том, что параметр прямого преобразования (открытый элемент) может быть в явном виде применен в системе. Параметр обратного преобразования (конфиденциальный элемент) хранится в секрете и доступен только одному абоненту (пользователю).

Однако, существует практическая сложность реализации обозначенных операций как программно, так и аппаратно. Ситуация усугубляется использованием в асимметричных криптографических системах классов вычетов больших целых чисел.

Криптографические алгоритмы основаны на математических вычислениях, стойкость асимметричных криптосистем основывается, как на сложности вычисления дискретного логарифма, так и на разложении числа на множители. Модульное умножение – наиболее часто встречающаяся операция при построении алгоритмов криптографии с открытыми ключами. С вычислениями по модулю удобно работать потому, что они ограничивают диапазон всех промежуточных величин и результатов. Для модуля N длиной k бит промежуточные результаты любого сложения, вычитания или умножения будут не длиннее $2k$ бит. Поэтому возведение в степень в модулярной арифметике можно выполнить без генерации очень больших промежуточных результатов.

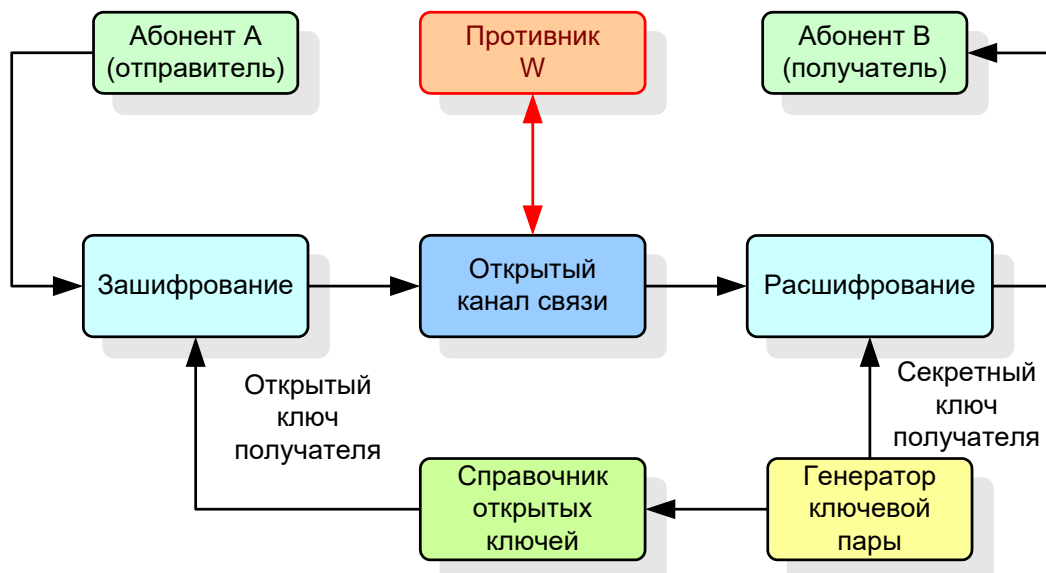


Рис. 1. Модель асимметричной криптосистемы

Поскольку многие криптографические алгоритмы основаны на возведении в степень по модулю, скорость вычисления операции взятия по модулю, будет вносить определяющий вклад в быстродействие алгоритма.

Существуют стандартные методы вычисления степени числа, a по модулю N : $a^x \bmod N$, где необходимо выполнить ряд умножений и делений и стандартные способы сделать это быстрее. Поскольку эти операции дистрибутивны, быстрее произвести возведение в степень как ряд последовательных умножений, выполняя каждый раз приведение по модулю.

Например, если необходимо вычислить: $a^4 \bmod N$, применяют не примитивный подход трех последовательных перемножений, а выполняют два малых умножения и два малых приведения по модулю: $(a^2 \bmod N)^2 \bmod N$, аналогично проводят вычисления для всех x , которые являются степенью числа 2.

Вычисление $a^x \bmod N$, где x – не является степенью числа 2. Проводится следующим образом, представляется число x как двоичная запись, что позволяет представить число x как сумму степеней 2.

Пусть $x = 13_{(10)} = 1101_{(2)}$, т.е. $13 = 2^3 + 2^2 + 2^0$.

Тогда

$$a^{13} \bmod N = ((a(a^{12})) \bmod N) = (a \times a^8 \times a^4) \bmod N = a \times ((a^2)^2)^2 \times (a^2)^2 \bmod N = (((a^2 \times a)^2)^2 \times a) \bmod N.$$

При накоплении промежуточных результатов потребуется только четыре умножения: $((((a^2 \bmod N) \times a \bmod N)^2 \times \bmod N)^2 \times a) \times \bmod N$.

Этот метод уменьшает трудоемкость вычислений $1,5 \times k$ операций в среднем, где k – длина числа в битах.

Если выполнять модульное умножение непосредственно, сначала перемножая целые числа (многочлены), затем, вычисляя остаток от деления, как описано выше, то сложность алгоритма будет определяться процедурой нахождения остатка. Использование математических преобразований на основе арифметики Монтгомери не требует операции деления, что позволяет значительно снизить его сложность. Выигрыш от использования арифметики Монтгомери будет заметен при работе с длинными числами, 200 бит и более, что и необходимо, например, при реализации цифровой подписи.

С использованием арифметики Монтгомери для вычисления произведения чисел A и B по модулю нечетного числа N , где A , B и N – большие числа, выбирается число R , равное степени числа 2 и превышающее N . N -вычет числа A по модулю N определяется как $AR \pmod N$, а произведение Монтгомери по модулю N как $MP(A, B, N, R) \equiv ABR^{-1} \pmod N$. В основу вычисления произведения Монтгомери положено преобразование Монтгомери, базирующееся на следующей теореме.

Теорема. Пусть N и R – взаимно простые положительные целые числа, $N' \equiv N^{-1}(\text{mod } R)$. Тогда для всех чисел T и $M \equiv TN'(\text{mod } R)$:

- число $\frac{T + MN}{R}$ – целое;

- $TR^{-1}(\text{mod } N) = \frac{T + MN}{R}$ или $TR^{-1}(\text{mod } N) = \frac{T + MN}{R} - N$.

Для вычисления произведения Монтгомери достаточно перемножить операнды как целые числа и затем выполнить преобразование Монтгомери. Для перевода числа в N -вычет нужно вычислить $\text{MP}(A, R^2(\text{mod } N), N, R)$. Для обратного преобразования N -вычета в число нужно вычислить произведение $\text{MP}(AR(\text{mod } N), 1, N, R)$.

Имеют место равенства:

$\text{MP}(AR(\text{mod } N), BR(\text{mod } N), N, R) \equiv ABR(\text{mod } N)$,

$AR + BR \equiv (A+B)R(\text{mod } N)$.

Таким образом, множество N -вычетов с операциями умножения по Монтгомери и сложения является кольцом. Нетрудно видеть, что это кольцо изоморфно кольцу классов вычетов по модулю N , изоморфизм задается умножением на R , обратный изоморфизм задается умножением на R^{-1} :

$\text{MP}(AR(\text{mod } N), R^{-1}R(\text{mod } N), N, R) \equiv AR^{-1}R(\text{mod } N) \equiv A(\text{mod } N)$.

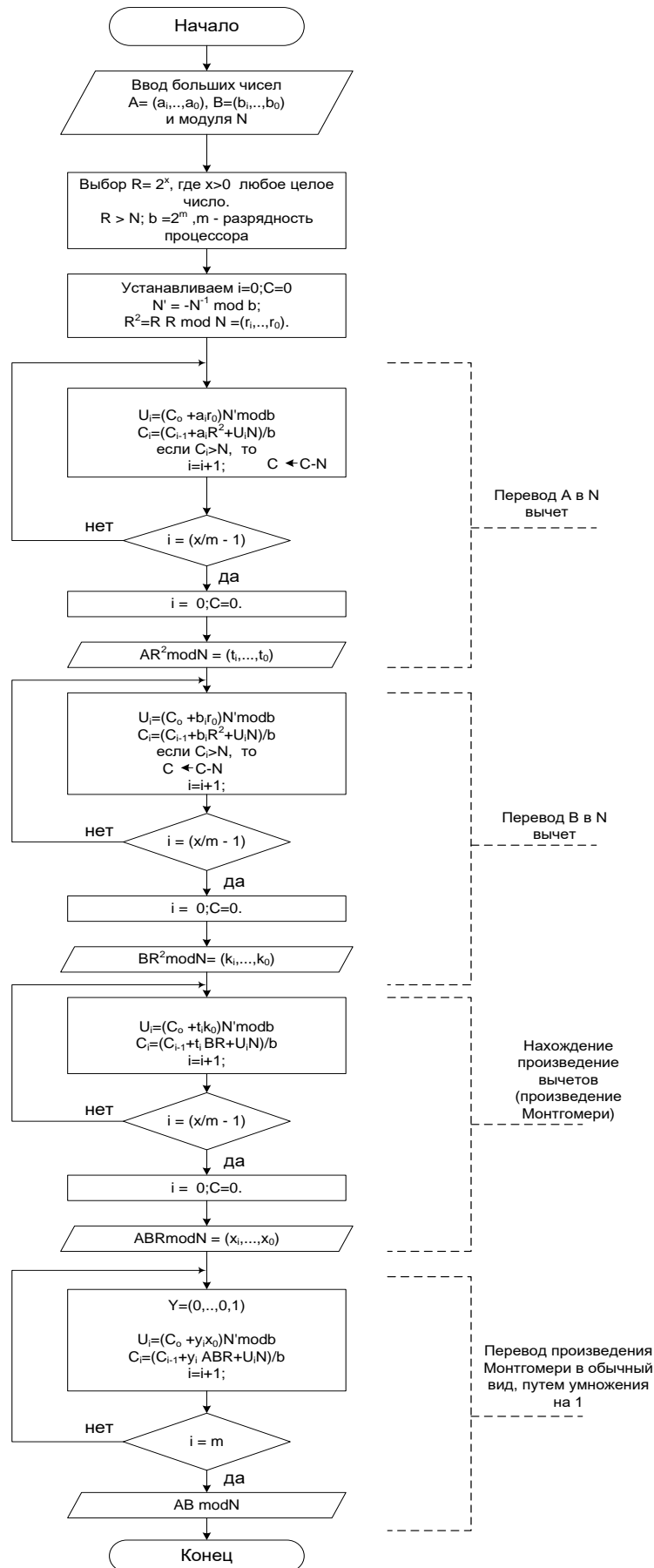
В арифметике Монтгомери можно выполнять не только умножение, но также сложение и возведение в степень. При этом нет необходимости переводить промежуточные результаты из N -вычетов в обычное представление. Как уже было сказано ранее, для выполнения арифметических операций методом Монтгомери необходимо использовать вспомогательные переменные N' , R . Причем при реализации данных операций с использованием аппаратно-ориентированных средств $N' = -N^{-1} \text{mod } b$, где N^{-1} – обратная величина, R – число равное степени числа 2 и превышающее N , $b=2^m$ где m – разрядность процессора (например, $m = 8$). Выбор оптимального способа нахождения вспомогательных переменных не вносит определяющего вклада в работу алгоритма в целом.

При работе с математическими преобразованиями методом Монтгомери можно условно выделить три этапа:

- нахождение вычетов сомножителей;
- вычисление операции по Монтгомери (произведение, сложение, возведение в степень);
- перевод в обычный вид.

а. Нахождение вычетов сомножителей

При умножении двух больших простых чисел A и B вычисление произведения Монтгомери $AR^2 \text{mod } N$, происходит следующим образом: предварительно вычисляется преобразование Монтгомери (U_i). $U_i = (C_0 + a_i r_i) N' \text{mod } b$, $C_i = (C_{i-1} + a_i R^2 + U_i N) / b$, если $C_i > N$, тогда C присваиваем разность $C = C - N$, где N' вспомогательная величина вычисленная выше, a_i – значение разряда сомножителя записанного в шестнадцатеричном формате, i – номер разряда (1,...,m), N – модуль по которому производим умножение. $R > N$, равно степени числа 2. Значение b описано выше. Причем операция взятия модуля по b , не подразумевает деления, просто берется последний байт числа, полученного в результате операций сложения и умножения. Деление на b при реализации аппаратно идентично сдвигу влево на восемь разрядов. Таким образом, данная методика позволяет избежать трудоемких операций взятия большого числа по большому модулю, путем сведения к вычислениям по малому модулю. Аналогично находится вычет для B . На первой итерации C_0 присваивается нулевое значение, затем значение нулевого разряда предыдущего значения. Количество итераций определяется используемой разрядностью. В результате имеем два вычета: числа A по модулю N (AR), и числа B по модулю N (BR).



**Рис. 2. Модульное умножение методом Монтгомери
б. Вычисление произведения Монтгомери $(AR)(BR) \bmod N$**

Вычеты представляются в виде $AR = (t_i, \dots, t_1)$, $BR = (k_i, \dots, k_1)$, дальше работа происходит аналогично вычислению вычетов, вычисляем преобразование и произведение Монтгомери: $U_i = (C_0 + t_i k_1)N \bmod b$, $C_i = (C_{i-1} + t_i BR + U_i N) / b$. Результат произведения вычетов представляется как $ABR = (x_i, \dots, x_1)$. Блок-схема модульного умножения методом Монтгомери приведена на рисунке 2.

с. Перевод произведения Монтгомери в обычный вид

Необходимо произвести умножение произведения вычетов ABR на единицу по Монтгомери. Операции производятся аналогично вышеописанным, и имеют вид: $U_i = (C_0 + y_i x_1)N \bmod b$, $C_i = (C_{i-1} + y_i ABR + U_i N) / b$, где y_i – значение разряда $Y = (0, \dots, 0, 1)$, причем i – номер разряда $(1, \dots, m)$. При переводе произведения Монтгомери в обычный вид получаем вычет произведения по модулю N .

В арифметике Монтгомери, как уже говорилось выше, можно выполнять не только умножение, но также сложение и возведение в степень, которые производятся аналогично:

- при сложении методом Монтгомери находится вычет для каждого суммируемого операнда t . Затем производится сложение вычетов и перевод результата в обычный вид, шаг 3;

- с помощью математических преобразований Монтгомери операция возведения в степень будет выглядеть аналогично операции умножения, за исключением лишь только того, что при возведении в степень; вычет числа будет производиться однократно (1 шаг). Вычисление произведения Монтгомери (2 шаг) будет производиться такое количество раз, чему равна сумма весов единиц. Если $n = 2^x$, то умножение производится x раз, где n – степень числа. Если x не степень числа два, то алгоритм разложения идентичен алгоритму описанному выше. Причем не стоит переводить промежуточные результаты из N вычетов в обычное представление, перевод в обычное представление осуществляется только при получении конечного результата (3 шаг).

Алгоритмы Монтгомери, как и стандартные алгоритмы взятия модуля, принадлежат к классу детерминированных алгоритмов, сложность которых оцениваются некоторым полиномом n . Для многочлена степени не более $2n-2$ при приведении по модулю $f(t) = t^n + t + b$ стандартным методом, требуется $n-1$ операции умножения и $2n-2$ операций сложения. При приведении по модулю $f(t)$ используя быструю арифметику Монтгомери учитывая, ненужность перевода промежуточных результатов из N -вычетов в обычный вид, то количество операций будет составлять $2n(n+1)$ обычных умножений слов.

Использование математического преобразования Монтгомери позволяет увеличить скорость работы криптографического алгоритма, как минимум в четыре раза, что расширяет область применения криптографии в телекоммуникационной сфере.

Асимптотическая сложность алгоритма умножения Монтгомери такая же, как и у алгоритма, выполняющего сначала умножение операндов «в столбик», а затем деление с остатком. Поэтому необходима более точная оценка сложности. На практике часто требуется выполнение не одиночного умножения, а цепочки умножений и сложений. В этом случае можно исключить преобразование произведения из N -вычета в обычную форму и преобразование операндов в N -вычеты. Тогда сложность умножения алгоритмом Монтгомери составляет $2n(n+1)$ обычных умножений слов.

Алгоритм Монтгомери используется для умножения в кольце классов вычетов $\mathbf{Z}/N\mathbf{Z}$. Однако вместо кольца целых чисел \mathbf{Z} можно использовать и произвольное евклидово кольцо. Например, если требуется вычислить произведение многочленов в кольце $k[x]/N(x)k[x]$, то выбирают вспомогательный многочлен $R(x)$, например, $R(x) = x^{n+1}$, ($\deg R(x) > \deg N(x)$), находят алгоритмом Евклида многочлены $R^{-1}(x)$, $N'(x)$ такие, что $R(x)R^{-1}(x) - N(x)N'(x) = 1$. Затем проводят такие же вычисления, как и в стандартном алгоритме. Отличие заключается в том, что на последнем шаге после деления на $R(x)$ всегда получается $\deg U(x) < \deg N(x)$, поэтому операция вычитания $N(x)$ не требуется.

Если арифметика Монтгомери используется для сложения точек эллиптической кривой и вычисления ведутся в проективной форме с тремя координатами XR , YR , ZR , то для перехода к аффинной форме нет необходимости сначала переходить от арифметики

Монтгомери к нормальному виду, так как при делении $x = \frac{XR}{ZR}$ происходит автоматическое сокращение на R .

Практическая программная реализация подтвердила осуществимость алгоритмов умножения Монтгомери, высокие характеристики быстродействия и приемлемую сложность их реализации. В настоящее время ведется интенсивная работа по совершенствованию алгоритмов быстрого выполнения модульных операций с целью увеличения быстродействия и упрощения его программных, а также аппаратных реализаций.

На рисунке 3 приведена качественная сравнительная диаграмма быстродействия работы алгоритмов, реализующих рассмотренные методы.

Авторами рассмотрен вопрос разработки быстрых алгоритмов выполнения модульных операций (умножения, сложения, возведение в степень) с использованием арифметики Монтгомери для выполнения указанных модульных операций в полях больших целых чисел в криптографических системах. Практическая программная реализация подтвердила осуществимость алгоритмов, высокие характеристики быстродействия и небольшую сложность их реализации.

В настоящее время ведется интенсивная работа по совершенствованию быстрых алгоритмов выполнения модульных операций с применением арифметики Монтгомери с целью увеличения быстродействия и упрощения их программных, а также аппаратных реализаций.

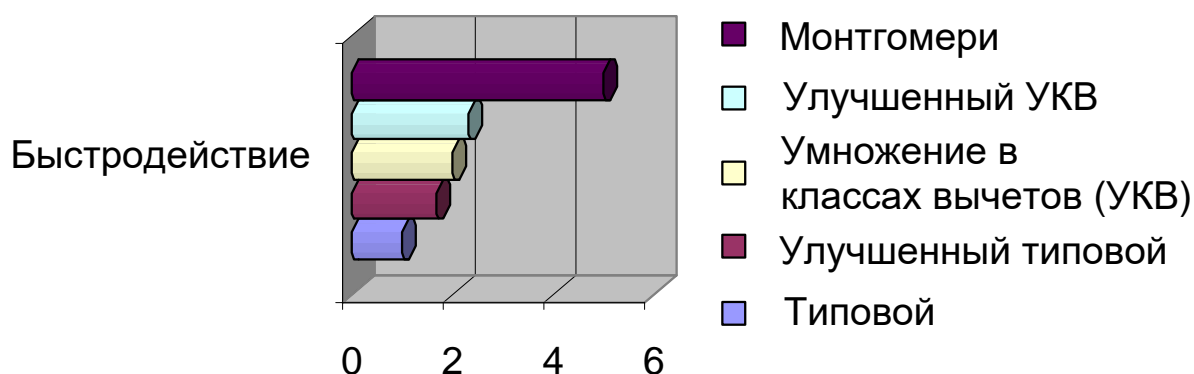


Рис. 3. Быстродействие работы различных алгоритмов

Библиографический список

1. Борисенков И.А., Дороненков М.Н., Николаев Д.Б., Мартынов А.П., Фомченко В.Н. Концепция информационного взаимодействия участников процесса разработки наукоемкой продукции. В сборнике: Информатизация образования-2014. 2014. С. 176-178.
2. Немченко И.А., Николаев Д.Б. Обеспечение безопасности качественной составляющей информации с использованием стохастических алгоритмов. Вестник Рязанского государственного университета имени С.А. Есенина. 2011. № 35. С. 126.
3. Мартынов А.П., Николаев Д.Б., Соколов С.Ю. Концептуальные основы построения систем обеспечения взаимной аутентификации объектов // Известия института инженерной физики, 2008. № 10(10). С.6-9 .
4. Костюков В.Е., Николаев Д.Б., Мартынов А.П., Фомченко В.Н. Динамическая аутентификация группы разнородных объектов. Информатизация образования. 2014. Т. 2014. С. 260.

5. Сплюхин Д.В., Николаев Д.Б. Анализ новейших требований ФСТЭК и общие решения существующих проблем защиты информационных систем. Сборник материалов X-ой Всероссийской молодежной научно-инновационной школы «Математика и математическое моделирование». 2016. С. 28-29.
6. Грибунин В.Г., Костюков В.Е., Мартынов А.П., Николаев Д.Б., Фомченко В.Н. Стеганографические системы. Критерии и методическое обеспечение: Учебно-методическое пособие // Под ред. В.Г. Грибунина. Саров: ФГУП «РФЯЦ-ВНИИЭФ», 2016, 324 с.
7. Шишкин Г., Николаев Д. Селекторы цифровых команд. Часть 1 // Компоненты и технологии. 2009. № 6 (95). С.96-101.
8. Пат. 2154722. Электронный кодовый замок / Мартынов А.П., Николаев Д.Б., Шишкин Г.И. Заявл. 27.08.1999, опубл. 20.08.2000. Бюл. № 23.
9. Шишкин Г., Николаев Д. Селекторы цифровых команд. Часть 2 // Компоненты и технологии. 2009. № 8 (97). С.112-116.
10. Пат. 2402810, 27.01.2009. Устройство защиты от несанкционированного доступа к информации / Курочкин А.А., Мартынов А.П., Николаев Д.Б., Фомченко В.Н. Патент на изобретение № 2402810 от 27.01.2009.
11. Грибунин В.Г., Костюков В.Е., Мартынов А.П., Николаев Д.Б., Фомченко В.Н. Стеганографические системы. Цифровые водяные знаки: Учебно-методическое пособие // Под ред. В.Г. Грибунина. Саров: ФГУП «РФЯЦ-ВНИИЭФ», 2016, 295 с.
12. Бабанов Н.Ю., Мартынов А.П., Николаев Д.Б., Фомченко В.Н., Новиков А.В. Виртуальная интерактивная система формирования и отработки управляющей информации. Вестник НГИЭИ. 2016. № 4 (59). С. 15-29.
13. Мартынов А.П., Николаев Д.Б., Мартынова И.А. Криптографические системы и метод факториального сжатия информации // Известия института инженерной физики, 2016. № 4(42). С.54-57.
14. Грибунин В.Г., Костюков В.Е., Мартынов А.П., Николаев Д.Б., Фомченко В.Н. Стеганографические системы. Атаки, пропускная способность каналов и оценка стойкости: Учебно-методическое пособие // Под ред. В.Г. Грибунина. Саров: ФГУП "РФЯЦ-ВНИИЭФ", 2015.
15. Пат. 2623894, 07.09.2017. Способ преобразования данных с равновероятностной инициализацией / Мартынов А.П., Мартынова И.А., Марунин М.В., Николаев Д.Б., Фомченко В.Н. Патент на изобретение № 2623894 от 07.09.2017.
16. Пат. 2408924, 10.01.2011. Устройство обеспечения безопасного формирования параметров / Курочкин А.А., Мартынов А.П., Николаев Д.Б., Фомченко В.Н. Патент на изобретение № 2408924 от 10.01.2011.
17. Пат. 2256053. Устройство управления электронными замками / Гончаров С.Н., Мартынов А.П., Николаев Д.Б., Фомченко В.Н. Заявл. 20.01.2003, опубл. 10.07.2005. Бюл. № 19.
19. Мартынов А.П., Волков К.О., Николаев Д.Б. Обеспечение безопасного взаимодействия компонентов интегрированной системы. Инновации в условиях развития информационно-коммуникационных технологий. 2008. № 1. С. 136-138.
20. Пат. 2497277, 27.10.2013. Способ сжатия двоичных данных в виде структурированных информационных блоков / Николаев Д.Б., Мартынов А.П. Патент на изобретение № 2497277 от 27.10.2013.
21. Пат. 2630429, 18.11.2016. Голихин М.В., Способ преобразования низкоэнтропийных сообщений / Мартынов А.П., Николаев Д.Б., Фомченко В.Н. Патент на изобретение № 2630429 от 18.11.2016.
22. Одинцов М.В., Сплюхин Д.В., Николаев Д.Б. Исследование вопросов оптимизации параметров защищенности информации. Сборник материалов IX-ой Всероссийской молодежной научно-инновационной школы «Математика и математическое моделирование». 2015. С. 75.

23. Пат. 2546238, 07.10.2013. Способ транслитерационного преобразования информации и передачи ее по каналам связи / Грибунин В.Г., Мартынов А.П., Николаев Д.Б., Фомченко В.Н. Патент на изобретение № 2546238 от 07.10.2013.
24. Пат. 2554525, 07.10.2013. Способ преобразования информации с синхронной сменой инициализирующих последовательностей в блоках, соединенных каналом связи с неопределенным периодом смены / Фомченко В.Н., Мартынов А.П., Николаев Д.Б., Масыгин А.М., Мартынов А.А. Патент на изобретение № 2554525 от 07.10.2013.
25. Пат. 2416169, 20.07.2009. Способ аутентификации объекта / Мартынов А.П., Николаев Д.Б. Патент на изобретение № 2416169 от 20.07.2009.
26. Пат. 2506633, 26.07.2012. Устройство хранения данных (варианты) / Фомченко В.Н., Николаев Д.Б., Гончаров С.Н., Жердин О.А., Силаев А.Г., Марунин М.В., Моксяков А.Н. Патент на изобретение № 2506633 от 26.07.2012.
27. Гончаров С., Николаев Д., Никитин В., Писецкий В. Схемотехническая реализация автомата. Компоненты и технологии. 2013. № 2 (139). С. 126-128.
28. Шишкин Г., Николаев Д. Селекторы цифровых команд. Часть 3 // Компоненты и технологии. 2009. № 9 (98). С.116-120.
29. Пат. 2224075. Электронно-механический замок / Мартынов А.П., Николаев Д.Б., Сычев В.Т., Шишкин Г.И. Заявл. 28.11.2001, опубл. 20.02.2004. Бюл. № 5.
30. Шишкин Г., Николаев Д. Селекторы цифровых команд. Часть 4 // Компоненты и технологии. 2009. № 11 (100). С.102-106.
31. Васильев Р.А., Николаев Д.Б. Анализ возможностей применения голосовой идентификации в системах разграничения доступа к информации. Научный результат. Информационные технологии. 2016. Т. 1. № 1. С. 48-57.
32. Астайкин А.И., Мартынов А.П., Николаев Д.Б., Фомченко В.Н. Методы и средства обеспечения программно-аппаратной защиты информации. Научно-техническое издание / Саров, 2015.
33. Пат. 2187612, 20.08.2002. Электронный кодовый замок / Мартынов А.П., Николаев Д.Б., Шишкин Г.И. Патент на изобретение № 2187612 от 08.05.2001.

S.V. Gruzdev, Y.V. Shilova K.D. Ermakov

**APPLICATION OF THE MONTGOMERY METHOD FOR ENGINEERING
APPLICATIONS OF PHYSICO-MATHEMATICAL SCIENCES**

Sarov Institute of Physics and Technology - branch of NRNU MEPhI

The issue of development of fast algorithms for performing modular operations (multiplication, addition, exponentiation) using Montgomery arithmetic to perform the indicated modular operations in fields of large integers in cryptographic systems is considered.

Keywords: performance, cryptographic systems, modular operations, software implementation, complexity.

ТЕОРЕТИЧЕСКИЕ И ПРАКТИЧЕСКИЕ АСПЕКТЫ ПОСТРОЕНИЯ ЭКСПЕРИМЕНТАЛЬНОЙ ПЛАТФОРМЫ ДЛЯ АПРОБАЦИИ ИНТЕЛЛЕКТУАЛЬНЫХ ИНФОРМАЦИОННЫХ ОБРАЗОВАТЕЛЬНЫХ СИСТЕМ

Саровский физико-технический институт – филиал НИЯУ МИФИ

Предложена интегрированная система сквозного проектирования для образовательных направлений, связанных с разработкой электронных устройств и приборов

Ключевые слова: интегрированная система, моделирование, надежность, система, проектирование, стойкость

Интеллектуальные информационные системы находят применение в самых разных отраслях, и в каждой прикладной области отмечаются специфические задачи, предполагающие формирование и верификацию исходных данных, измерение функциональных возможностей, определение показателей и критериев качества прикладных интеллектуальных технологий. Еще более сложные задачи возникают при создании и эксплуатации интеллектуальных информационных систем в области образования, поскольку они должны учитывать не только техническую и технологическую основу конкретной области знаний, но и обеспечивать формирование заданного уровня подготовки операторов автоматизированных систем «человек-машина». Подготовка полноценных специалистов невозможна без обеспечения практических занятий и моделирования ситуационных кейсов, решение которых максимально приближено к реальным условиям, в которых действуют потребители интеллектуальных технологий. Второй задачей при создании и эксплуатации интеллектуальных информационных образовательных систем является их адаптация к требуемому информационному контенту в рамках осуществления образовательной деятельности. Подобная адаптация невозможна без применения современных средств имитационного моделирования в совокупности с гибко настраиваемыми аппаратными структурами (экспериментальной платформы для апробации интеллектуальных информационных образовательных систем).

С теоретической точки зрения построение экспериментальной платформы включает в себя несколько этапов:

- разработка и анализ имитационной модели процессов образовательной деятельности с целью определения исходных данных и функциональных требований к информационной системе;
- синтез модели структуры взаимодействия информационной системы;
- исследование парка используемых программно-аппаратных средств;
- формирование функциональных блоков экспериментальной платформы по направлениям образовательной деятельности;
- определение базового состава программных образовательных продуктов и формирование рабочих мест с использованием функциональных блоков программного обеспечения;
- разработка алгоритмов информационного взаимодействия участников образовательного процесса.

Таким образом, при построении экспериментальной платформы используется функционально-ориентированная структура, поскольку, решающую роль в организации образовательного процесса играет четко определенная иерархическая функциональная структура, связанная с образовательными программами и соответствующими

профессиональными стандартами. При этом функционально-ориентированные структуры обладают рядом недостатков, затрудняющих внедрению и развитию современных интеллектуальных технологий.

Недостатки функционально-ориентированной иерархии существенно затрудняют выделение и формализацию процессов, протекающих при осуществлении образовательной деятельности. Однако для построения имитационной модели процессов образовательной деятельности может быть применен один из современных инструментов аналитики – процессный подход. Применение процессного подхода позволяет отразить самые важные аспекты образовательной деятельности, целостно показывающие, как происходит достижение компетентностных целей, поставленных перед обучаемыми. В качестве аппарата функционального моделирования образовательной деятельности можно применить методологию ARIS. Следует отметить, что для моделирования процесса образования описанная методология применена впервые. Ранее она использовалась для моделирования деятельности предприятий банковской сферы, сферы товаров и услуг.

Новизна предложенного процессного подхода заключается в разработке математической модели функционирования информационной образовательной системы и реализующей ее модели управления процессом обучения на ее основе. Данные модели позволяют определить характерные особенности каждого компонента информационной системы в области осуществления информационной деятельности, информационного взаимодействия, моделирования изучаемого материала, автоматизации процессов управления. Практически эти модели реализуются в экспериментальной платформе в виде подпрограмм взаимодействия с информационной образовательной системой. В частности, эти подпрограммы могут оценивать и влиять на такие аспекты средств информационно-компьютерных технологий как методические, организационные, материально-технические и профилактические мероприятия с позиций теории обучения, в том числе выявлении особенностей обеспечения дистанционного обучения.

Экспериментальная платформа, взаимодействующая с информационной образовательной системой, дает возможность пользователю извлекать и реконструировать важное для него знание, выполнять различные логические преобразования информации, осуществлять поиск удобных способов ее кодирования, обработки и т.п. Может быть реализована пошаговая отработка учебного процесса и реализующих его методик с выполнением аналитических срезов закрепления материала или моделирования ситуаций степенной градаций освоения знаний. Экспериментальная платформа, по сути является не автономной системой преобразования информации, а своеобразным информационным партнером человека. Технический аспект взаимодействия непосредственно определяется программно-аппаратной составляющей экспериментальной платформы как машинным средством хранения и преобразования информации, информационный аспект связан с тем ансамблем ролей, которые реализует обучаемый, используя средства информационно-компьютерных технологий для образовательных и профессиональных целей; коммуникация не сводится к передаче информации, а информирование – к аккумуляции информации. Осуществление этих процессов опирается на эвристические алгоритмы, заложенные в экспериментальной платформе и уровень знаний пользователя. Экспериментальной платформой апробируются также и функции автоматизированного контроля, реализуемого в информационной образовательной системе, в условиях создания сложных алгоритмов анализа ответов обучаемых, и наличия обратной связи.

Таким образом, теоретический аспект построения экспериментальной платформы для апробации интеллектуальных информационных образовательных систем сводится к анализу структуры информационной образовательной системы и математических моделей информационных процессов, протекающих в ней в условиях реализации теории обучения, исследованию основных характеристик логической структуры и содержания учебного материала, структуры и математической модели информационной образовательной системы. После проведенного анализа формируется модель модели управления процессом

обучения в информационной образовательной системе с позиции теории управления сложным объектом (основа экспериментальной платформы), определяются особенности и функции автоматизированного контроля, реализуемого в условиях создания сложных алгоритмов анализа ответов обучаемых; в выявлении содержательных и методических аспектов подготовки преподавателей технических вузов.

Практическая реализация экспериментальной платформы осуществляется посредством процедурных модулей, управляемых системой «Управление проектом», целью которой является обеспечение возможности групповой работы в процессе образовательной деятельности, то есть просмотра в реальном времени и совместного использования фрагментов общих информационных ресурсов. При этом одни фрагменты данных могут обновляться регулярно, а другая часть информации остается статичной. Такой распределенный доступ к спискам материалов и конфигурации образовательных программ облегчает процедуру и резко сокращает время на доступ к информационным ресурсам, экономя, тем самым, временные, материальные и трудовые ресурсы [13-16].

а) управление составом учебных программ.

Подсистема «Управление проектом» содержит информацию о составе учебных программ, их версиях и конфигурациях. Важной особенностью является наличие нескольких представлений состава учебных программ для различных предметных областей (начальный, средний, продвинутый уровень и т.д.), а также управление внутренней структурой учебной программы.

б) классификация.

Подсистема «Управление проектом» позволяет производить распределение учебных программ и методических документов в соответствии с различными классификаторами. Это может быть использовано при автоматизации поиска образовательных дисциплин с нужными компетенциями для формирования учебных модулей и блоков или для работы в автоматическом режиме с последовательным изучением материала [20-23].

в) календарное планирование и выполнение учебных процессов.

Подсистема «Управление проектом» содержит функции формирования календарного плана работ, распределения ресурсов по отдельным задачам и контроля выполнения задач со стороны руководства.

г) вспомогательные функции, обеспечивающие взаимодействие подсистемы «Управление проектом» с другими программными средствами, с пользователями, а также взаимодействие пользователей друг с другом.

Система «Управление проектом» неразрывно связано с хранилищем данных экспериментальной платформы, которое представляет собой распределенную базу данных по приборам, основанную на системе управления базами данных, реализованной на основе Microsoft SQL Server и предназначенную для хранения и управления всей информацией, необходимой на всех этапах учебного процесса.

Местоположение системы «Управление проектом» в общей структурной схеме экспериментальной платформы и взаимодействие всех компонентов экспериментальной платформы приведены на рис. 1.



Рис. 1. Местоположение системы «Управление проектом» в структурной схеме экспериментальной платформы

Серверная часть системы «Управление проектом» включает в себя следующие модули:

- 1) модуль администрирования, предназначенный для управления серверами баз данных, серверами приложений и собственно базами данных подсистемы «Управление проектом» и всей экспериментальной платформы;
- 2) модуль конфигурации, служащий для настройки баз данных экспериментальной платформы;
- 3) клиентский модуль подсистемы «Управление проектом»;
- 4) модуль формирования отчетов, предназначенный для автоматизированного формирования отчетов в удобном для просмотра и печати виде об объектах базы данных экспериментальной платформы;
- 5) модуль автоматизации учебных процессов WorkFlow, предназначенный для сопровождения и организации рабочих учебных процессов;
- 6) модуль справочников, предназначенный для хранения необходимой информации экспериментальной платформы.

В результате работы рассмотрены теоретические и практические аспекты построения экспериментальной платформы для апробации интеллектуальных информационных образовательных систем и осуществлена ее программно-аппаратная реализация.

Библиографический список

1. Немченко И.А., Николаев Д.Б. Обеспечение безопасности качественной составляющей информации с использованием стохастических алгоритмов. Вестник Рязанского государственного университета имени С.А. Есенина. 2011. № 35. С. 126.
2. Шишкин Г., Николаев Д. Селекторы цифровых команд. Часть 2 // Компоненты и технологии. 2009. № 8 (97). С.112-116.
3. Васильев Р.А., Николаев Д.Б. Анализ возможностей применения голосовой идентификации в системах разграничения доступа к информации. Научный результат. Информационные технологии. 2016. Т. 1. № 1. С. 48-57.
4. Грибунин В.Г., Костюков В.Е., Мартынов А.П., Николаев Д.Б., Фомченко В.Н. Стеганографические системы. Цифровые водяные знаки: Учебно-методическое пособие // Под ред. В.Г. Грибунина. Саров: ФГУП "РФЯЦ-ВНИИЭФ", 2016. 295 с.
5. Пат. 2546238, 07.10.2013. Способ транслитерационного преобразования информации и передачи ее по каналам связи / Грибунин В.Г., Мартынов А.П., Николаев Д.Б., Фомченко В.Н. Патент на изобретение № 2546238 от 07.10.2013
6. Шишкин Г., Николаев Д. Селекторы цифровых команд. Часть 1 // Компоненты и технологии. 2009. № 6 (95). С.96-101.
7. Пат. 2256053. Устройство управления электронными замками / Гончаров С.Н., Мартынов А.П., Николаев Д.Б., Фомченко В.Н. Заявл. 20.01.2003, опубл. 10.07.2005. Бюл. № 19.
8. Пат. 2154722. Электронный кодовый замок / Мартынов А.П., Николаев Д.Б., Шишкин Г.И. Заявл. 27.08.1999, опубл. 20.08.2000. Бюл. № 23.
9. Шишкин Г., Николаев Д. Селекторы цифровых команд. Часть 3 // Компоненты и технологии. 2009. № 9 (98). С.116-120.
10. Пат. 2546238, 07.10.2013. Способ транслитерационного преобразования информации и передачи ее по каналам связи / Грибунин В.Г., Мартынов А.П., Николаев Д.Б., Фомченко В.Н. Патент на изобретение № 2546238 от 07.10.2013.
11. Шишкин Г., Николаев Д. Селекторы цифровых команд. Часть 4 // Компоненты и технологии. 2009. № 11 (100). С.102-106.
12. Бабанов Н.Ю., Мартынов А.П., Николаев Д.Б., Новиков А.В., Фомченко В.Н. Виртуальная интерактивная система формирования и отработки управляющей информации. Научный журнал «Вестник НГИЭИ», № 4 (59), 2016 г., стр. 15-29.
13. Пат. 2554525, 07.10. 2013.Способ преобразования информации с синхронной сменой инициализирующих последовательностей в блоках, соединенных каналом связи с неопределенным периодом смены / Фомченко В.Н., Мартынов А.П., Николаев Д.Б., Мясягин А.М., Мартынов А.А. Патент на изобретение № 2554525 от 07.10.2013
14. Одинцов М.В., Сплюхин Д.В., Николаев Д.Б. Исследование вопросов оптимизации параметров защищенности информации. Сборник материалов IX-ой Всероссийской молодежной научно-инновационной школы «Математика и математическое моделирование». 2015. С. 75.
15. Пат. 2416169, 20.07.2009. Способ аутентификации объекта / Мартынов А.П., Николаев Д.Б. Патент на изобретение № 2416169 от 20.07.2009.
16. Пат. 2224075. Электронно-механический замок / Мартынов А.П., Николаев Д.Б., Сычев В.Т., Шишкин Г.И. Заявл. 28.11.2001, опубл. 20.02.2004. Бюл. № 5.
17. Пат. 2623894, 07.09.2017. Способ преобразования данных с равновероятностной инициализацией / Мартынов А.П., Мартынова И.А., Марунин М.В., Николаев Д.Б., Фомченко В.Н. Патент на изобретение № 2623894 от 07.09.2017.
18. Пат. 2408924, 10.01.2011. Устройство обеспечения безопасного формирования параметров / Курочкин А.А., Мартынов А.П., Николаев Д.Б., Фомченко В.Н. Патент на изобретение № 2408924 от 10.01.2011.

19. Пат. 2506633, 26.07.2012. Устройство хранения данных (варианты) / Фомченко В.Н., Николаев Д.Б., Гончаров С.Н., Жердин О.А., Силаев А.Г., Марунин М.В., Моксяков А.Н. Патент на изобретение № 2506633 от 26.07.2012.
20. Пат. 2402810, 27.01.2009. Устройство защиты от несанкционированного доступа к информации / Курочкин А.А., Мартынов А.П., Николаев Д.Б., Фомченко В.Н. Патент на изобретение № 2402810 от 27.01.2009.
21. Пат. 2497277, 27.10.2013. Способ сжатия двоичных данных в виде структурированных информационных блоков / Николаев Д.Б., Мартынов А.П. Патент на изобретение № 2497277 от 27.10.2013 г.
22. Пат. 2630429, 18.11.2016. Голихин М.В., Способ преобразования низкоэнтропийных сообщений / Мартынов А.П., Николаев Д.Б., Фомченко В.Н. Патент на изобретение № 2630429 от 18.11.2016

V.A. Aleshkin, K.D. Ermakov

**THEORETICAL AND PRACTICAL ASPECTS OF BUILDING AN EXPERIMENTAL
PLATFORM FOR TESTING INTELLIGENT INFORMATION EDUCATIONAL
SYSTEMS**

Sarov Institute of Physics and Technology - branch of NRNU MEPhI

An integrated system of end-to-end design for educational areas related to the development of electronic devices and devices is proposed

Keywords: integrated system, modeling, reliability, system, design, durability

ОБЗОР АКТУАЛЬНЫХ ВОПРОСОВ БИОМЕТРИЧЕСКОЙ ИДЕНТИФИКАЦИИ НА ОСНОВЕ ОСОБЕННОСТЕЙ КЛАВИАТУРНОГО ПОЧЕРКА

Нижегородский государственный технический университет им. Р.Е. Алексеева

Защита информации и управление доступом к информационным ресурсам часто обеспечиваются системами идентификации пользователей. Настоящая статья посвящена клавиатурному почерку как средству идентификации. Статья включает обзор исследований применения особенностей клавиатурного почерка для идентификации пользователей при вводе произвольного текста и постоянного контроля личности пользователя при работе с компьютером.

Ключевые слова: биометрия, биометрическая идентификация, клавиатурный почерк, управление доступом, информационная безопасность.

Введение

Вопросы безопасности компьютерных систем, управления доступом к информационным ресурсам и защиты «оцифрованной» частной жизни от неправомерного использования давно вошли в повседневный дискурс. К этому привели сразу несколько значимых факторов: растущая информатизация, экспансия информационных технологий, повышение компьютерной грамотности и, как следствие, развитие новых методов несанкционированного доступа к информации.

В настоящее время парольная защита является наиболее популярной мерой решения обозначенных вопросов. Вместе с тем применение паролей несет целый ряд рисков. Пароли могут быть подобраны, утеряны, украдены, подсмотрены в процессе ввода, перехвачены и т.д. В некоторой степени часть рисков снимается за счет внедрения двухфакторной аутентификации, когда кроме пароля для получения доступа требуется дополнительное подтверждение при помощи кода, имеющего ограниченное время действия. Однако и здесь нельзя говорить об обеспечении должной степени безопасности на протяжении всего времени взаимодействия пользователя с компьютерной системой.

Вариантом решения упомянутых проблем является использование клавиатурного почерка. Данный метод является одним из наиболее доступных способов биометрической идентификации, так как реализуется с помощью программных средств с применением стандартного оборудования – устройств ввода. Кроме доступности стоит отметить не менее важную характеристику данного подхода с точки зрения решения задач информационной безопасности. Такой метод предполагает, что процесс идентификации является для пользователя неявным и может осуществляться непрерывно. То есть в отличие от стандартной парольной защиты, о которой пользователю известно заранее, можно осуществлять дополнительный контроль без уведомления пользователя и быть уверенным в личности пользователя не только на этапе ввода пароля, но и далее, когда начать работу с компьютером может любое другое лицо. Так же применение методов анализа клавиатурного почерка может быть полезным при реализации сегментированного доступа разных категорий пользователей к информационным ресурсам, например, появляется возможность контролировать использование платных интернет-сервисов пожилыми людьми или детьми.

Анализ источников и литературы

История публикаций о применении анализа особенностей клавиатурного почерка началась со статьи «Клавиатурное устройство для идентификации личности» в IBM Technica Disclosure Bulletin в 1975 году [1], где Р. Спиллейн впервые рассмотрел возможность идентификации на основе особенностей клавиатурного набора текста.

Развитие исследований в данном направлении продолжилось в вооруженных силах США. Группа исследователей под руководством Форзена в Римской Лаборатории ВВС США опубликовала статью об изучении более, чем 30 физиологических особенностей пользователей, которые могут быть обнаружены при анализе клавиатурного почерка [2]. Затем в 1980-м Гейнс и др. [3] опубликовали результаты применения динамики нажатия клавиш для аутентификации пользователей на основе t-критерия Стьюдента.

Лишь в середине 90-х появляется интерес к непрерывной идентификации – С. Шеферд публикует материал о «непрерывной аутентификации путем анализа характеристик набора текста с клавиатуры». В 1997 году Монроуз и Рубин [4] применяют анализ клавиатурного почерка для идентификации на основе произвольного текста. В их исследовании эффективность идентификации на основе клавиатурного набора произвольного текста оказалась лишь на уровне 23%, что с самого начала показало высокую сложность разработки систем непрерывной идентификации.

Популярность этого подхода в научной среде возросла в начале 2000-х годов, когда стал отмечаться значительный рост числа публикаций об исследованиях применения технологии распознавания особенностей клавиатурного почерка. Большое число работ в этом направлении опубликовано в первые десятилетия XXI века: изучение особенностей нажатия клавиш и отдельных характеристик клавиатурного почерка получило свое развитие в многочисленных работах отечественных [5,6,7,8,9,10] и зарубежных ученых [11,12,13,14].

Уровень точности и надежность результатов работы описываемых систем идентификации не всегда отвечает предъявляемым требованиям в рамках решения прикладных задач, что не позволяет начать их промышленную эксплуатацию.

Целью статьи является анализ перспектив идентификации пользователей на основе клавиатурного почерка для повышения безопасности и надежности систем идентификации при работе с компьютером.

Актуальное положение и перспективы

Можно выделить два основных применения биометрических систем идентификации пользователей на основе особенностей ввода текста с клавиатуры. В первую очередь, это дополнительная проверка в момент авторизации пользователя. В этом случае авторизация пользователя при помощи ввода логина и пароля должно подвергаться дополнительному контролю за счет анализа особенностей клавиатурного почерка. Во-вторых, после авторизации, на протяжении всего времени работы, становится доступной непрерывная проверка личности пользователя на основе его клавиатурного почерка.

Таким образом, системы идентификации, основанные на клавиатурном почерке, могут функционировать на двух уровнях: статическом (fixed-text), анализирующем фиксированный, заранее известный системе текст, и динамическом (free-text), обрабатывающем произвольный, заранее не известный системе текст. И в том, и в другом случае реализация системы предполагает подготовительный этап – обучение системы, в ходе которого происходит сбор образцов клавиатурного почерка пользователей, и непосредственно этап идентификации, когда система сравнивает особенности клавиатурного почерка конкретного пользователя с образцами, собранными на первом этапе.

Обучение системы для динамического распознавания клавиатурного почерка при вводе произвольного текста ведется в течение нескольких дней путем сбор образцов и мониторинга повседневного взаимодействия пользователя с клавиатурой. После сбора образцов и формирования профилей (наборов отличительных характеристик клавиатурного почерка) создается классификатор, на основе которого будут выявляться сходства и различия между профилями.

Основной механизм работы системы состоит в сравнении имеющихся в базе эталонных профилей с данными, получаемыми с клавиатуры. Для этого применяется широкий набор методов: от простых статистических средств анализа до более сложных моделей распознавания образов и алгоритмов нейронных сетей.

Отдельно необходимо отметить, что при применении всех систем биометрической идентификации существует широкий разброс точности и риск ошибок идентификации. Такие ошибки и не всегда достаточный уровень точности детерминированы рядом причин, связанных не только с реализацией системы распознавания, но и особенностями самого пользователя, и внешних свойств среды, в которой находится пользователь. Так различные результаты идентификации могут быть получены при использовании разных устройств (компьютеров и клавиатур), при наборе знакомого и незнакомого пользователю текста, при разной длине текстов и т.д. Навык набора текста, особенно десятипальцевым методом, также играет существенную роль, так как опытные пользователи обладают большей стабильностью при наборе текста.

Ошибки можно разделить на два класса: система может ошибочно идентифицировать постороннего пользователя как доверенного и разрешить доступ, или некорректно ограничить доступ доверенному пользователю в следствие неверной идентификации. Эти особенности использования биометрической идентификации требуют выработки критериев, определяющих надежность системы. Эти критерии возможно определить через рассчитываемые показатели эффективности: % False Accept Rate – доля ошибочно предоставленных доступов, % False Reject Rate – доля ошибочных отказов в доступе, Time Of Lockout – время между первым нажатием клавиши и ограничением доступа, то есть скорость распознавания постороннего пользователя [15,16].

При этом стоит отметить широкие перспективы использования систем биометрической идентификации по клавиатурному почерку. В дополнение к непрерывному контролю личности пользователя, работающего с компьютерными системами в рамках обеспечения информационной безопасности, можно применять системы на основе анализа особенностей клавиатурного почерка для обнаружения вторжений, проведения криминологических исследований и криминалистических экспертиз [17, с. 98] путем создания профилей правонарушителей и преступников, оценки текущего физического и психо-эмоционального состояния, когда, например, пользователи работают с потенциально опасными компьютерными системами или системами жизнеобеспечения (АЭС, лечебные учреждения, службы спасения и т.д.).

Выводы

Системы биометрической идентификации, основанные на анализе особенностей клавиатурного почерка, обладают целым рядом преимуществ: могут использоваться незаметно для пользователя на основе их повседневной работы с клавиатурой, они не требуют затрат на закупку и внедрение дополнительного оборудования. В то же время недостаточная надежность и нестабильная эффективность, связанная с внешними факторами и особенностями самих пользователей, не позволяет применять такие системы самостоятельно, а только в качестве компонента сложных комплексов обеспечения информационной безопасности. При этом возможна реализация как систем с распознаванием фиксированного текста, так и произвольного. И вторые имеют значительное более широкий спектр применения. Дополнительно стоит заметить, что применение систем биометрической идентификации на основе клавиатурного почерка возможно не только при работе с классической компьютерной клавиатурой, но и на мобильных устройствах, банкоматах, сейфах и любых других устройствах с клавиатурными системами ввода.

Библиографический список

1. Spillane R., «Keyboard Apparatus for Personal Identification», IBM Technical Disclosure Bulletin, vol. 17, no. 3346, 1975.
2. Forsen G., Nelson M., and Staron R. Jr. «Personal attributes authentication techniques», Technical Report RADC-TR-77-333, Rome Air Development Center, October 1977.
3. Gaines R., Lisowski W., Press S., and Shapiro N., «Authentication by keystroke timing: some preliminary results», Rand Rep. R-2560-NSF, Rand Corporation, 1980.
4. Monroe F. and Rubin A., «Authentication via Keystroke Dynamics» / Proceedings of the Fourth ACM Conference on Computer and Communication Security, Apr 1 1997. – Zurich, Switzerland, pp. 48-56.
5. Крутохвостов Д.С., Хиценко В.Е. Парольная и непрерывная аутентификация по клавиатурному почерку средствами математической статистики. / Вопросы кибербезопасности. – М.: АО «НПО «Эшелон», 2017. – №5(24). – С. 91-99.
6. Еременко А. В., Левитская Е. А., Сулавко А. Е., Самотуга А. Е. Разграничение доступа к информации на основе скрытого мониторинга пользователей компьютерных систем: непрерывная идентификация / Вестник СибАДИ. – Выпуск 6 (40). 2014.– С. 82-102.
7. Брюхомицкий Ю.А. Цепочный метод клавиатурного мониторинга // «Известия ЮФУ. Технические науки. Информационная безопасность. 2009. – №11. – С. 135-145.
8. Брюхомицкий Ю.А. Клавиатурный мониторинг на основе иммунологического клонирования / Безопасность информационных технологий. 2016 г. – № 4. – С. 5-11.
9. Казарин Н. М. Разработка и исследование методов скрытого клавиатурного мониторинга: автореф. дис. на соискание науч. степени канд. техн. наук: спец. 05.13.19 / 2006. – 20 с.
10. Абашин В. Г. Автоматизация процесса определения психофизического состояния оператора автоматизированного рабочего места в АСУТП: автореф. дис. на соискание науч. степени канд. техн. наук: спец. 05.13.06 / 2008. – 18 с.
11. Davoudi H. and Kabir E., «A new distance measure for free text keystroke authentication» / Proceedings 14th International CSI Computer Conference (CSICC '09), Oct. 2009, pp. 570–575.
12. Messerman A., Mustafic T., Camtepe S. A., and Albayrak S., «Continuous and non-intrusive identity verification in real time environments based on free-text keystroke dynamics» / Proceedings International Conference on Biometrics (IJB '11), October 2011, pp. 1–8.
13. Gunetti D. and Picardi C., «Keystroke analysis of free text» / ACM Transactions on Information and System Security, vol. 8, no. 3, 2005, pp. 312–347.
14. Shimshon T., Moskovitch R., Rokach L., and Elovici Y., «Continuous verification using keystroke dynamics» / Proceedings of the International Conference on Computational Intelligence and Security (CIS'10), December 2010, pp. 411–415.
15. Bours P., «Continuous Keystroke Dynamics: a Different Perspective Towards Biometric Evaluation» / Information Security Technical Report, Vol. 17, 2012, pp. 36-43.
16. Bours P. and Barghouthi H., «Continuous Authentication Using Biometric Keystroke Dynamics» / Proceedings of the Norwegian Information Security Conference, 2009, pp. 41-58.
17. Ahmed A., Traore I., Almulhem A. «Digital Fingerprinting Based on Keystroke Dynamics» / Proceedings International Symposium on Human Aspects of Information Security & Assurance (HAISA), 2008, pp. 94-104.

P.N. Mashtanov, M.V. Martynyuk

**REVIEW OF CURRENT ISSUES OF BIOMETRIC IDENTIFICATION BASED ON
THE FEATURES OF KEYSTROKE DYNAMICS**

Nizhny Novgorod state technical university n.a. R.E. Alekseev

The paper introduces an approach to the protection of access for computer systems and confidential information based on keystroke dynamics. This article reviews the sequence of research into the issue. An identification system using keystroke dynamics is considered to be an additional component of the information security complex. Additionally, the article describes the perspectives and scope of application of this method, its advantages, and drawbacks.

Keywords: biometric, biometric authentication, keystroke dynamics, authentication, information security

ПРИМЕНЕНИЕ СИСТЕМЫ RSA В ЗАДАЧАХ КРИПТОГРАФИИ И ИССЛЕДОВАНИЕ ЕЕ КРИПТОСТОЙКОСТИ

Нижегородский государственный технический университет им. Р.Е. Алексеева

В работе рассматривается криптографический алгоритм шифрования RSA и математические методы, на которых он построен; его использование в задачах криптографии, в том числе для электронной цифровой подписи; анализируется криптостойкость RSA путем применения некоторых видов атак. По результатам проведенных исследований сделан вывод о криптостойкости RSA, сформулированы рекомендации по выбору параметров алгоритма шифрования.

Ключевые слова: асимметричная криптосистема, алгоритм шифрования, RSA, криптостойкость, атака повторным шифрованием, атака Винера, китайская теорема об остатках, бесключевое чтение, электронная цифровая подпись.

Введение

С появлением телекоммуникаций и электронного документооборота появилась необходимость в надежной передаче сообщений, а также сохранении их целостности и подлинности. В связи с этим в последние десятилетия особенно бурно развиваются криптографические системы защиты информации, основанные на разных принципах действия. Первой криптосистемой с открытым ключом является система RSA (аббревиатура от фамилий разработчиков - Rivest, Shamir, Adleman), используемая как для шифрования/дешифрования информации, так и для электронной цифровой подписи (ЭЦП). Настоящая работа посвящена исследованию использования системы RSA в задачах криптографии и ее криптостойкости.

Особенности реализации криптосистемы RSA

Криптосистема RSA реализована на принципе однонаправленных функций с модификацией и основана на вычислительной сложности операции факторизации больших целых чисел. Надежность криптосистемы строится на следующих математических принципах:

– При заданных параметрах p и q достаточно легко вычислить модуль $n = pq$, однако задача разложения числа n на простые множители считается сложной.

– При известных числах x , e , n достаточно легко вычислить $y = x^e \bmod n$, но вычисление значения x при известных числах y , e , n является трудной задачей. Для получения значения x необходимо знать разложение $n = pq$ на простые множители и d такое, что $de \equiv 1 \bmod \varphi(n)$, тогда - $x = y^d \bmod n$.

Алгоритм RSA состоит из четырех этапов: генерация ключей; распределение ключей; шифрование и дешифрование сообщений с помощью сгенерированных ключей.

Для реализации криптосистемы RSA, необходимы алгоритмы, обеспечивающие точное вычисление криптографических параметров, в частности необходима генерация случайных простых чисел p и q . Эффективный алгоритм генерации простых случайных чисел должен соответствовать двум требованиям:

– иметь близкое к равномерному распределение сгенерированных простых чисел на множестве всех k -битных простых чисел;

– не иметь возможности воспроизведения процесса генерации конкретного случайного простого числа при использовании криптостойкого генератора псевдослучайных чисел.

Для обеспечения безопасности передачи данных криптосистемой RSA при выборе параметров рекомендуется соблюдать следующие условия [1]:

- Размер простых чисел p и q должен быть не менее 512 бит, размер модуля n – не менее 1024 бит.
- Числа $(p - 1)$ и $(q - 1)$ должны иметь по крайней мере один большой простой множитель.
- Простые числа p и q должны быть выбраны независимо друг от друга. Если разность между p и q слишком мала, то их можно вычислить с использованием метода факторизации Ферма:

$$\left(\frac{p+q}{2}\right)^2 - n = \left(\frac{p-q}{2}\right)^2. \quad 1)$$

- Простые числа p и q должны быть уникальными для каждого пользователя, иначе, если два пользователя имеют одно из двух общих простых чисел, можно вычислить модуль с помощью алгоритма Евклида и расшифровать сообщения одного из пользователей.
- Для противодействия атаке Хастада число пользователей в системе должно быть меньше, чем значение e .
- Для противодействия атаке Франклина-Рейтера значение e должно быть большим (рекомендуется использовать $e = 2^{16} + 1 = 65537$).
- Для противодействия атаке Винера нужно использовать показатель d не меньший, чем $n^{0,292}$.
- Короткие сообщения должны быть дополнены с помощью технологии оптимального асимметричного дополнения шифрования (OAEP) [2].

Исходя из представленных условий, можно сделать вывод, что криптоскопкость асимметричного алгоритма RSA напрямую зависит от размера ключей; кроме этого, для обеспечения надежной работы алгоритма необходимо обеспечить высокую скорость выполнения операций шифрования/дешифрования при использовании больших целых чисел. Однако вычислительные машины не имеют стандартных типов данных необходимой разрядности. Поэтому при реализации RSA используются алгоритмы длинной арифметики для выполнения арифметических операций с большими целыми числами, такие как [4]:

- длинное умножение («в столбик», методом Карацубы-Оффмана, методом Шенхаге-Штрассена, с помощью быстрого преобразования Фурье) – для возведения в степень числа;
- длинное деление («в столбик», с помощью схемы Горнера) – для взятия остатка от деления;
- сложение, вычитание и сравнение длинных чисел – для работы алгоритма деления длинных чисел;
- умножение длинного числа на короткое – для деления длинных чисел.

Применение алгоритма RSA в задачах криптографии

В настоящее время алгоритм RSA применяется как в виде самостоятельных криптографических средств, так и в качестве встроенных средств в значительном количестве криптографических приложений, включая PGP, S/MIME, TLS/SSL, IPSEC/IKE и пр., кроме того RSA используется в сочетании с симметричными алгоритмами (AES, Serpent, Twofish и др.). Отсюда следует, что криптосистему RSA можно считать наиболее распространенной среди других асимметричных алгоритмов.

Использование алгоритма RSA в сочетании с симметричными алгоритмами связано с тем, что последние выполняют шифрование больших объемов данных с более высокой скоростью, поэтому часто используют смешанный алгоритм шифрования, при котором создается случайный сеансовый ключ, а затем с помощью полученного ключа и симметричных алгоритмов происходит шифрование и дешифрование сообщений; с

помощью алгоритма RSA – шифрование и дешифрование только сеансового ключа. Сеансовый ключ существует только в рамках передачи конкретного сообщения.

Кроме задач шифрования/дешифрования алгоритм RSA широко используется и для создания электронной цифровой подписи. ЭЦП – это данные в цифровой форме, которые добавляются к другим данным в цифровой форме (подписываемой информации), и используются для определения лица, подписывающего информацию [4].

Основой электронной цифровой подписи является математическое преобразование подписываемых данных с использованием личного закрытого ключа подписывающего и выполнением следующих условий:

- Цифровая подпись создается только с помощью использования личного закрытого ключа автора документа.
- Проверку подлинности электронной цифровой подписи может осуществить любой человек, имеющий доступ к соответствующему открытому ключу.
- Электронную подпись делает недействительной любое, даже малейшее изменение подписанного документа, а именно, если изменится всего лишь один бит в документе огромного размера.

Необходимо отметить, что, как правило, передаваемый документ не шифруется и в исходном виде передается получателю. Однако, применяя алгоритм RSA, возможно создавать документы, которые будут и зашифрованы, и подписаны цифровой подписью.

Результатом работы стало компьютерное приложение по созданию и проверке ЭЦП на основе алгоритма RSA. Для разработки приложения использовались программная среда Microsoft Visual Studio 2015 и язык программирования C#. Интерфейс программного приложения представлен на рис. 1.

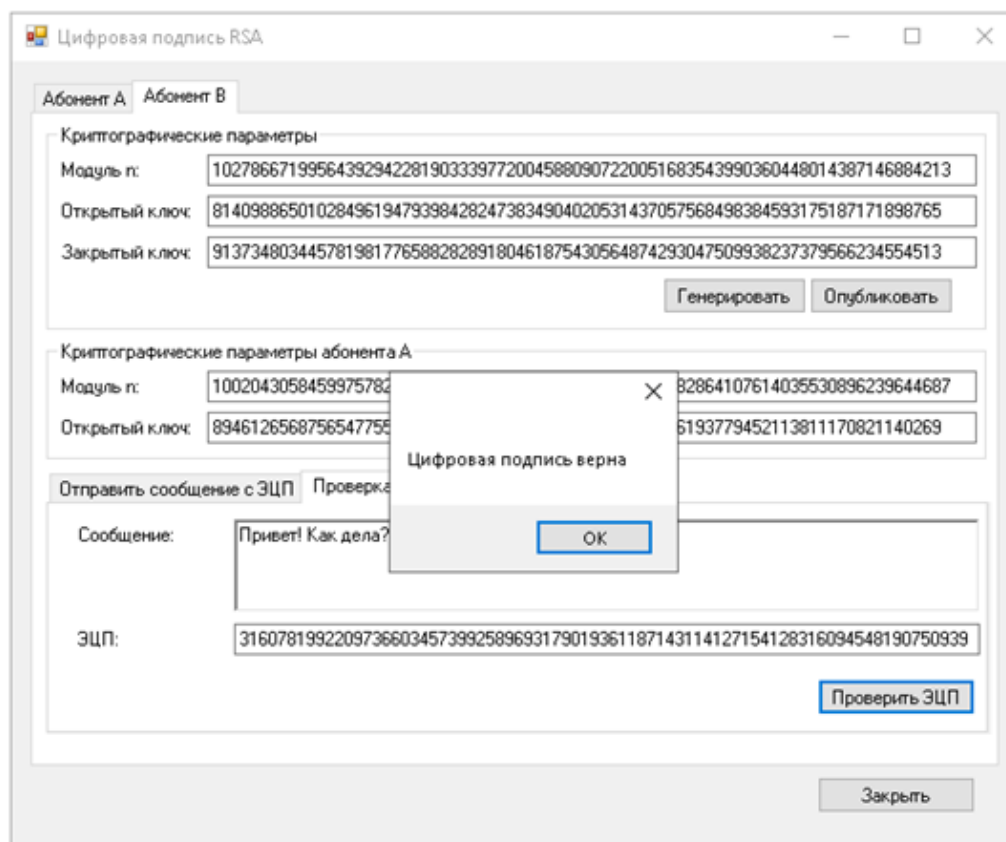


Рис. 1. Интерфейс работы приложения

Приложение имеет следующие функциональные возможности:

- Генерация ключей и вычисление модуля Абонента А и Абонента В.
- Публикация (обмен) открытых ключей и модулей обоих абонентов.

- Написание сообщения одним из абонентов для другого абонента.
- Вычисление ЭЦП для сообщения.
- Отправка сообщения с электронной подписью.
- Проверка ЭЦП сообщения.

Анализ криптостойкости системы RSA

Одной из задач работы было исследование криптостойкости алгоритма RSA. Для этой цели были выбраны и программно-реализованы следующие атаки: атака методом повторного шифрования, атака Винера, атака на основе китайской теоремы об остатках и атака методом бесключевого чтения; которые помогли выявить, при какой ошибке в выборе параметров алгоритм станет небезопасным. Программное приложение реализовано на ОС Linux с использованием программной среды для разработки приложений IntelliJIDEA 2018.3 Community Edition.

Разработанные алгоритмы атак дополнены возможностью измерения времени работы (в мс) для дальнейшего анализа их быстродействия. Для расчета времени работы алгоритма берется системное время перед началом атаки и после завершения расшифровки полученные данные записываются в файл. Для большей достоверности вычисляется среднее время расшифрования, которое рассчитывается за 50 повторов взлома одного шифра. Такой подход позволил максимально снизить уровень влияния аппаратуры на время работы алгоритма атаки на шифр RSA.

В работе принято считать «маленьким сообщением» шифр длиной 219 символов с учетом пробелов, а «большим сообщением» - шифр, содержащий 971 символ. Для их успешного взлома исходные шифры преобразуются в соответствии с требованиями каждого метода:

- Для метода бесключевого чтения нужно иметь два сообщения с общим модулем.
- Для метода на основе китайской теоремы об остатках необходимо иметь три шифра, и обеспечить значение открытой экспоненты e равное 3.
- Для метода, основанного на теореме Винера, должно быть малое значение секретной экспоненты.
- Для метода повторного шифрования нет особых ограничений, поэтому в файл будет записан шифр с соответствующей парой открытого ключа.

Таким образом, будет создано по четыре файла с шифром каждого сообщения.

В ходе работы были проведены вычислительные эксперименты, в результате которых исходные шифры были взломаны. В результате сравнения были вычислены средние значения времени работы каждого алгоритма (табл. 1):

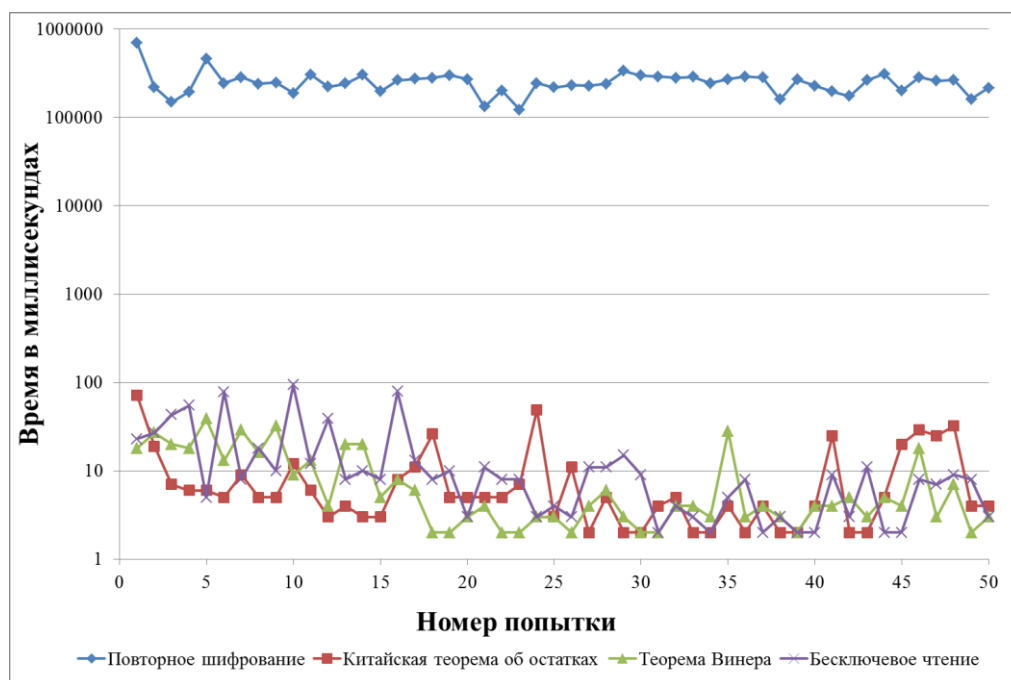
Таблица 1

Среднее время работы алгоритмов в миллисекундах

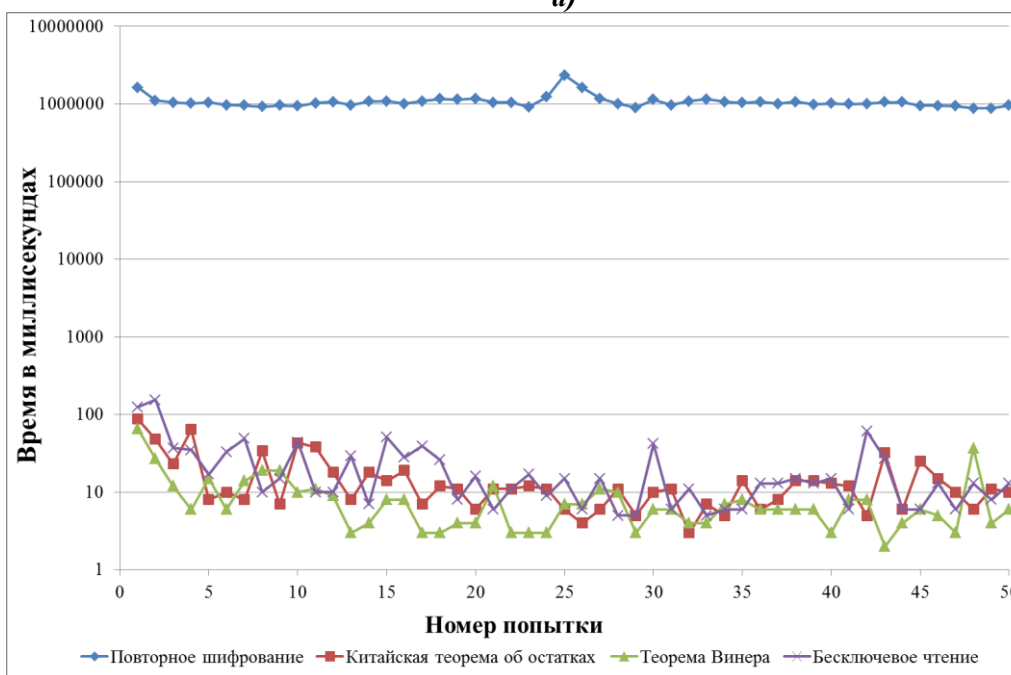
Количество символов в шифре, шт	Метод повторного шифрования, мс	Китайская теорема об остатках, мс	Теорема Винера, мс	Метод бесключевого чтения, мс
219	255 720,9	9,78	8,92	14,58
971	1 077 344,0	15,94	9,08	22,7

Анализ показал, что время взлома методом повторного шифрования в разы больше, чем в трех других методах, которые затрачивают примерно одинаковое время при работе с шифрами одинаковой длины. В нашем случае время работы алгоритма зависит от метода атаки и объема шифруемых данных.

Представим результаты работы реализованных на языке программирования Java алгоритмов атак при малой длине зашифрованного сообщения (рис. 2, а), и при использовании большого количества зашифрованных данных (рис. 2, б).



а)



б)

Рис. 2. Результат работы программы
а) при малом количестве данных; б) при большом количестве данных

Как видно на графиках, в результате работы, время взлома шифра с использованием алгоритма повторного шифрования в несколько раз больше остальных, также на него сильно влияет длина шифра. Напротив, на методы, основанные на китайской теореме об остатках, теореме Винера и бесключевом чтении, длина шифра не сильно загружает процесс и время увеличивается несущественно. Погрешность в графиках обуславливается аппаратными характеристиками компьютера, на котором проходило тестирование.

Метод повторного шифрования затрачивает много времени, в нашем случае в среднем от 5 до 18 минут для малых и больших текстов соответственно. Однако этот алгоритм может расшифровать любой файл с шифрами, в то время как остальные методы могут расшифровать только файлы определенного (подготовленного) вида.

Анализ эффективности исследуемых атак

Проанализируем успешность проведенных атак.

Атака методом бесключевого чтения стала возможна, потому что в реализации алгоритма RSA ключи p и q не генерировались случайно, а дублировались. Как следствие – ключи были успешно вычислены с помощью алгоритма Евклида. Если p и q будут малы, то за счет этого увеличится длина d . Тогда эта уязвимость будет использована менее затратной атакой – методом повторного шифрования.

Для параметра e самым частым и рекомендованным значением является значение 65 537 – большое простое число Ферма, которое также гарантирует быстрое действие алгоритма RSA. Но если e будет повторяющимся малым значением, это приведет к возможному использованию атаки на основе китайской теоремы об остатках. При увеличении e значение d будет уменьшаться. Тогда метод повторного шифрования будет иметь меньший успех. Но, если значение секретной экспоненты станет ниже границы - $\sqrt[3]{\frac{n}{3}}$, то шифр будет взломан атакой на основе теоремы Винера.

Практика подтвердила, что нужно придерживаться рекомендаций к выбору любого из параметров и исключить частое повторение ключей.

Заключение

По результатам проведенной работы были сделаны следующие выводы:

- Стойкость криптосистемы RSA зависит от сложности задачи факторизации, поэтому наибольшую угрозу представляют атаки, основанные на решении этой задачи.
- Остальные атаки направлены на недостатки и особенности реализации криптосистемы, поэтому достаточно подобрать правильные параметры системы.
- Для обеспечения точности результатов работы алгоритма RSA и увеличения скорости его работы важен выбор математических методов, при котором учитываются их трудоемкость и сложность реализации.
- В ходе исследования при оценке стойкости алгоритма RSA программное приложение справилось со всеми шифрами, созданными с определенными уязвимостями за короткое время (табл.1):
 - a) Методом бесключевого чтения была получена расшифровка сообщений, т.к. использовался общий модуль при шифровании одного сообщения для двух абонентов.
 - b) Метод на основе китайской теоремы взломал шифры, потому что была выбрана открытая экспонента e равна 3.
 - c) Методу Винера удалось расшифровать сообщения в связи с использованием малой секретной экспоненты.
 - d) Метод повторного шифрованием справился с шифрами за самое большое количество времени, это стало возможным, т.к. использовалась длина ключа равная 64 бит.
- При соблюдении всех указаний по подбору параметров, исследованные атаки будут либо бесполезны, либо полученная информация будет уже не актуальна.

Дальнейшая работа будет направлена на изучение симметричных криптографических систем, в частности планируется исследовать работу симметричного блочного алгоритма «Кузнечик» (ГОСТ 34.12-2015), применение его в задачах криптографии, а также его криптостойкость с целью обнаружения уязвимостей и сравнения его эффективности с алгоритмом RSA.

Библиографический список

1. Маховенко Е.Б. Теоретико-числовые методы в криптографии: Учебное пособие. – М.: Гелиос АРВ, 2006. – 320 с., ил.
2. Оптимальное асимметричное шифрование с дополнением: [Электронный ресурс] // Википедия. – Режим доступа: [https://ru.wikipedia.org/wiki/Оптимальное асимметричное шифрование с дополнением](https://ru.wikipedia.org/wiki/Оптимальное_асимметричное_шифрование_с_дополнением), свободный. – Загл. с экрана (дата обращения: 10.05.2020).
3. Крэндэлл Ричард, Померанс Карл. Простые числа: Криптографические и вычислительные аспекты / Под ред. и с предисл. В. Н. Чубарикова. – М.: УРСС: Книжный Дом «ЛИБРОКОМ», 2011. – 664 с.
4. Федеральный закон от 6 апреля 2011 г. N 63-ФЗ "Об электронной подписи".
5. Ротков Л.Ю., Зобнев А.В. Электронная цифровая подпись в электронном документообороте – Нижний Новгород, 2006, 42с.
6. Хеш-алгоритмы: [Электронный ресурс] // Хабр. – Режим доступа: <https://habr.com/ru/post/93226/>, свободный. – Загл. с экрана (дата обращения: 15.04.2020).

О.Р. Timofeeva, А.А. Limarenko, А.V. Usanova, N.A. Farafonova

APPLICATION OF THE RSA SYSTEM IN CRYPTOGRAPHY PROBLEMS AND RESEARCH OF ITS CRYPTIC RESISTANCE

Nizhny Novgorod state technical university n.a. R.E. Alekseev

The paper considers the RSA cryptographic encryption algorithm and the mathematical methods on which it is built; its use in cryptography tasks, including for electronic digital signature; the cryptographic strength of RSA is analyzed by using some types of attacks. Based on the results of the research, a conclusion was made about the cryptographic strength of RSA, recommendations were formulated for choosing the parameters of the encryption algorithm.

Keywords: cryptosystem, encryption algorithm, RSA, cryptographic strength, re-encryption attack, Wiener attack, Chinese theorem on residuals, keyless reading, electronic digital signature, electronic digital signature RSA.

Е.М. Купаев¹, Ю.С. Федосенко¹, Н.П. Ямпурин², В.И. Логинов¹

ПОВЫШЕНИЕ КРИПТОСТОЙКОСТИ АЛГОРИТМОВ ШИФРОВАНИЯ НА ОСНОВЕ ИДЕИ БАТТИСТА ПОРТА

¹Волжский государственный университет водного транспорта

²Нижегородский технический университет им. Р.Е. Алексеева
г. Нижний Новгород

Рассмотрено использование идеи Баттиста Порты повышения криптостойкости простых алгоритмов замены и алгоритмов шифровальных машин симметричного шифрования. Предлагается заменить коллективный подход к шифрованию информации в современных каналах связи на индивидуальное шифрование.

Ключевые слова: шифры замены, симметричное шифрование, алгоритм Цезаря, Джамбаттиста делла Порты, коллективное шифрование, индивидуальное шифрование

Симметричное шифрование – это такие алгоритмы, которые используют один и тот же ключ для шифрования открытого текста и расшифровки шифротекста. Ключи могут быть как идентичными, так и преобразованием между двумя ключами. Обычно, ключ представляет собой некий секрет, который двое или более лиц могут использовать чтобы вести зашифрованный диалог. Основная черта таких алгоритмов является и их недостатком по сравнению с другим видом шифрования.

Краткая историческая справка: Джамбаттиста делла Порты (итал. Giambattista della Porta; 1.11.1535 – 4.02.1615 — итальянский врач, философ, алхимик и драматург. В части источников именуется Джованни Баттиста (делла) Порты (итал. Giovanni Battista della Porta) (далее Баттиста Порты). В 1563 опубликовал фундаментальный труд по криптологии — De furtivis literarum notis (Про скрытую значимость отдельных букв), сделавший его известным [1, 6].

Идея Баттиста Порты состоит в том, что для повышения криптостойкости алгоритмов шифрования необходимо выполнять замену алфавита, используя ключ, который в идеале должен быть равен длине используемого алфавита. При использовании короткого ключа при выполнении частотного анализа, незамаскированная ключом часть алфавита не будет подвержена замене и поэтому в этом случае возможна дешифровка ключа последующего алгоритма шифрования, например, шифра Цезаря.

Использование и реализация идеи Баттиста Порты позволяет решить некоторые задачи современной криптографии – повысить криптостойкость, увеличить эффективность и производительность методов шифрования, а также адаптировать классические методы шифрования для использования с современной компьютерной техникой. Кроме того, позволяет более эффективно и безопасно использовать шифровальные машины [3-4] и их компьютерные эмуляторы [5].

У алгоритмов шифрования замены существует уязвимость к частотному (статистическому) анализу, особенно уязвим к нему классический алгоритм шифрования Цезаря.

Цель работы – модификация алгоритма Цезаря для увеличения его криптостойкости на основе подхода Баттиста Порты, заключающаяся в манипуляции с алфавитом исходного текста.

Для модификации алгоритма сначала заменим алфавит исходного текста, совершив его перестановку, а затем к тексту с модифицированным алфавитом применим простейший алгоритм шифрования Цезаря [2]. Проведем частотный анализ и покажем

невозможность нахождения ключа с помощью статистического метода анализа шифротекста.

Авторами был разработан программный пакет для осуществить шифрование Цезаря, в том числе и его модификация. Такой пакет выполняет чтение файла с текстом, производит замену алфавита текста, затем выполняет шифрование Цезаря, и после этого записывает в новый файл зашифрованный текст. Далее проводится частотный анализ двух вариантов кодированного текста с предварительной заменой исходного алфавита по идее Баттиста Порты и без нее.

Поскольку сами эти алгоритмы обладают недостаточной криптостойкостью, то применим их комбинацию, таким образом сделав конечный шифр более криптостойким. Например, сначала применяем идею перестановки алфавита, а затем алгоритм замены, таким образом усложняя задачу взлома, потому что в данном случае нужно будет знать два ключа и порядок применения алгоритмов.

Рассмотрим пример использования алгоритма Цезаря без использования идеи Баттиста Порты для шифрования длинного текста (Том 1, Война и мир, Л.Н. Толстого). Проведем частотный анализ исходного текста и зашифрованного по алгоритму Цезаря с ключом $K=7$. Результаты частотного анализа и взлома ключа приведены в таблице. 1.

Таблица 1

Частотный анализ алгоритма Цезаря с ключом $K=7$

Исходный текст			Шифрование Цезаря, ключ 7			
Код символа	Количество	Вероятность	Код символа	Количество	Вероятность	Сдвиг
32	103411	20%	39	103411	20%	7
238	42745	8%	245	42745	8%	7
224	34286	6%	231	34286	6%	7
229	32473	6%	236	32473	6%	7
232	28149	5%	239	28149	5%	7
237	26132	5%	244	26132	5%	7
235	22473	4%	242	22473	4%	7
242	21828	4%	249	21828	4%	7
241	19969	4%	248	19969	4%	7
240	19440	4%	247	19440	4%	7
226	17454	3%	233	17454	3%	7
234	13791	3%	241	13791	3%	7
...						
...						
Всего:	527644	100%	Всего:	527644	100%	

Таким образом, из таблицы видно, что алгоритм отработал правильно, так как сдвиг каждого символа равен 7, как и было задано. Для взлома данного шифра можно предположить, что наиболее встречающиеся в тексте символы: пробел, буква «о» и буква «а». После определения того, что три символа сдвинуты на 7, полагаем что ключ к шифру Цезаря $K=7$, поэтому классический алгоритм шифрования Цезаря уязвим к простейшему частотному анализу.

Рассмотрим идею Баттиста Порты подмены исходного алфавита. В качестве ключа замены алфавита (ключ Баттиста Порты) можно использовать различные подходы. Например, формировать ключ длиной исходного алфавита, сгенерированный из последовательности неповторяющихся символов известного текста (подход предлагаемый Баттиста Порты). При использовании компьютерной эмуляции удобнее использовать

генератор псевдослучайных чисел для выполнения перестановки исходного алфавита, при этом длина ключа равна длине используемого алфавита (ASCII-8).

В нашем случае в качестве примера использован стандартный генератор псевдослучайных чисел современного компьютера, в качестве ключа использовано произвольное целое число (наш ключ Баттиста Порты), которое использовано для начальной инициализации генератора псевдослучайных чисел. Результаты кодирования и частотного анализа приведены в таблице 2.

Таблица 2

Частотный анализ модифицированного алгоритма Цезаря с заменой алфавита с использованием идеи Баттиста Порты

Исходный текст			Модифицированный шифр Цезаря			
Код символа	Количество	Вероятность	Код символа	Количество	Вероятность	Сдвиг
32	103411	20%	167	103411	20%	135
238	42745	8%	250	42745	8%	12
224	34286	6%	253	34286	6%	29
229	32473	6%	33	32473	6%	60
232	28149	5%	130	28149	5%	154
237	26132	5%	132	26132	5%	151
235	22473	4%	246	22473	4%	11
242	21828	4%	163	21828	4%	177
241	19969	4%	134	19969	4%	149
240	19440	4%	252	19440	4%	12
226	17454	3%	179	17454	3%	209
234	13791	3%	245	13791	3%	11
...						
...						
Всего:	527644	100%	Всего:	527644	100	

В результате анализа можно видеть, что сдвиг каждого символа практически не повторяется. В конкретном случае числа 7 нет в наиболее встречающихся символах. Таким образом, даже предположив наиболее частые символы в тексте и найдя их сдвиги, при взломе невозможно определить ключ, который использовался в алгоритме шифрования Цезаря.

Использование подхода с заменой алфавита по идее Баттиста Порты для других алгоритмов замены и перестановки позволит однозначно повысить криптостойкость шифрования, приближаясь по криптостойкости к современным компьютерно-ориентированным методам шифрования, которые в основе предлагают многоступенчатое кодирование по известным схемам (сеть Фейстеля). Многоступенчатый подход к шифрованию можно с успехом использовать для классических алгоритмов шифрования в связке с подходом Баттиста Порты.

Встроенные коллективные алгоритмы шифрования при высокой криптостойкости используемых алгоритмов обладают одним серьезным недостатком – ключи в этих системах шифрования назначаются автоматически и могут быть доступны большому числу людей, работникам спецслужб, что фактически для заинтересованных людей не является секретной информацией (вспомните случай с Цукербергом).

Кроме того, известно, что при шифровании с использованием алгоритмов с открытым ключом высока вероятность подмены ключа и, как следствие, получение недостоверной искаженной информации. Выход из этого положения возможен с использованием дополнительного индивидуального шифрования открытого ключа. При

этом алгоритм шифрования должен быть быстродействующим и его криптостойкость должна быть не менее 1-2 минут. Например, алгоритм Цезаря с модификацией Баттиста Порты вполне с этим справится.

Выводы

1. Приведенный шифр показывает, что использование простейших систем шифрования позволяет повысить их быстродействие, криптостойкость и открывает возможность выполнять дополнительное шифрование информации в современных телекоммуникационных системах индивидуально и независимо от встроенных (коллективных, корпоративных) алгоритмов шифрования.

2. Развитие индивидуальных систем шифрования в эпоху компьютерных технологий и телекоммуникаций требует адаптации классических алгоритмов шифрования под требования современных телекоммуникационных систем, расширения используемых алфавитов и (или) создания конверторов алфавитов с возможностью подмены алфавитов с использованием подхода на основе идеи Баттиста Порты, что позволяет многократно увеличить криптостойкость шифрования даже алгоритмов с низкой криптостойкостью.

Библиографический список

1. Джамбаттиста делла Порты – электронный ресурс. Режим доступа: https://ru.wikipedia.org/wiki/Джамбаттиста_делла_Порты
2. Гребенников В. В. Криптология и секретная связь. Сделано в СССР /В.В. Гребенников. – М.: Алгоритм, 2017. - 480 с.
3. Фиалка-125 – электронный ресурс. Режим доступа https://ru.wikipedia.org/wiki/Фиалка_М-125
4. Шифровальная машина Энигма – электронный ресурс. Режим доступа: <https://ru.wikipedia.org/wiki/Энигма>
5. Эмулятор шифровальной машины Энигмы – электронный ресурс. Режим доступа: <http://enigmaco.de/enigma/enigma.html>
6. Сمارт Н. Криптография. Москва: Техносфера, 2005. 528 с.

E.M. Kupaev¹, Ju.S. Fedosenko¹, N.P. Jampurin², V.I. Loginov¹

INCREASING THE CRYPTOGRAPHIC STRENGTH OF ENCRYPTION ALGORITHMS BASED ON THE IDEA OF BATTISTA PORTA

¹ *Volga State University of Water Transport, N. Novgorod*

² *Nizhny Novgorod state technical university n. a. R. E. Alekseev*

The use of Battista's idea to increase the cryptographic strength of simple substitution algorithms and algorithms of symmetric encryption machines is considered. It is proposed to replace the collective approach to information encryption in modern communication channels with individual encryption.

Keywords: substitution ciphers, symmetric encryption, Caesar's algorithm, Giambattista della Porta, collective encryption, individual encryption

МЕТОДЫ ЗАМЕНЫ ПРИ КОДИРОВАНИИ ИНФОРМАЦИИ С ПОДМЕНОЙ АЛФАВИТА

Волжский государственный университет водного транспорта

Представлены результаты разработки новых методов кодирования информации для аппаратуры передачи конфиденциальных данных

Ключевые слова: шифрование, кодирование

В настоящее время остро стоит проблема конфиденциальности разного рода информации, размещенной в базах данных и информационных системах государственных структур и частных предприятиях. Такие системы ныне используются практически во всех сферах жизни, что естественным образом породило необходимость защиты конфиденциальной информации от несанкционированного использования и кражи. В этих целях применяется шифрование.

Шифроваться могут не только текстовые данные, но и почти все форматы файлов, используемых операционными системами, т.е. разработаны и нашли свое практическое применение алгоритмы, способные шифровать любые типы цифровых данных, включая текстовые сообщения, исполняемые файлы, изображения и цифровые объекты другого вида.

Ко всем вышеперечисленным данным применимы и простые алгоритмы шифрования, реализованные в виде моделирующих программ. Надежность таких алгоритмов чрезвычайно мала.

Появление алгоритмов сквозного шифрования [1-3] и необходимость шифрования в реальном времени предъявляют к таким алгоритмам также и требования быстродействия, что делает использование популярных инструментов шифров нерациональным.

Соответственно, перед специалистами в области кибербезопасности неизменно стоит задача разработки алгоритмов шифрования, обеспечивающих наряду с высокой криптостойкостью шифров, также и достаточно высокое быстродействие с целью интеграции этих инструментов в программные приложения и службы.

Цель данной работы состоит в повышении криптостойкости симметричных шифров замены и расширения их алфавитов.

Основной идеей данной работы является использование алгоритма Баттисты Порты (далее БП) для подмены алфавита с длиной ключа равной длине алфавита.

Алгоритм БП представляет собой шифр замены с подменой алфавита с ключевым словом равным длине алфавита.

Для практического использования разрабатываемых алгоритмов замены с модификацией БП в информационных системах они должны быть адаптированы к работе с алфавитом ASCII. В этом случае общее число перестановок символов представляется числом $8,578 \cdot 10^{506}$, что позволяет приблизительно оценить время полного перебора в часах величиной $(08,578 \cdot 10^{506})/2$. Ключом шифра является число инициализации генератора случайных чисел BIOS ЭВМ. В нашем случае генератор случайных чисел управляется параметром типа данных Double.

Диапазон данных типа Double представляется собой интервал $[1.7e-308, 1.7e+308]$, что позволяет полностью перекрыть требования алгоритмов шифрования. Однако, реальное значение максимального числа равно 2^{64} , что соответствует длительности полного перебора 213503,9 часа при условии продолжительности синтеза варианта алфавита, не превышающего 1 нс.

Сводные данные о сдвигах символов при кодировании представлены в таблице 1. Первое число – код символа, затем символ и количество повторений символа в эталонном тексте.

Таблица 1

Сводные данные сдвигов символов шифра Цезаря

Эталонный текст	По алгоритму Цезаря код 3	Сдвиги символов алгоритму Цезаря	Сдвиги символов по алгоритму Цезаря+БП
227(г)= 251	228(д)= 240	227- 228= 1	227- 66= 161
252(ь)= 251	230(ж)= 251	252- 230= 22	252- 221= 31
231(з)= 257	234(к)= 257	231- 234= 3	231- 247= 16
251(ы)= 268	254(ю)= 268	251- 254= 3	251- 210= 41
239(п)= 306	242(т)= 306	239- 242= 3	239- 180= 59
243(у)= 334	246(ц)= 334	243- 246= 3	243- 162= 81
255(я)= 335	3()= 335	255- 3= 252	255- 127= 128
228(д)= 344	231(з)= 344	228- 231= 3	228- 79= 149
101(е)= 360	104(н)= 360	101- 104= 3	101- 173= 72
236(м)= 398	239(п)= 398	236- 239= 3	236- 199= 37
44(,)= 404	47(/)= 404	44- 47= 3	44- 228= 184
234(к)= 479	237(н)= 479	234- 237= 3	234- 164= 70
240(р)= 564	243(у)= 564	240- 243= 3	240- 212= 28
226(в)= 659	229(е)= 659	226- 229= 3	226- 200= 26
235(л)= 676	238(о)= 676	235- 238= 3	235- 207= 28
241(с)= 684	244(ф)= 684	241- 244= 3	241- 152= 89
242(т)= 755	245(х)= 755	242- 245= 3	242- 255= 13
232(и)= 883	235(л)= 883	232- 235= 3	232- 57= 175
237(н)= 1054	240(р)= 1054	237- 240= 3	237- 105= 132
229(е)= 1155	232(и)= 1155	229- 232= 3	229- 253= 24
224(а)= 1181	227(г)= 1181	224- 227= 3	224- 205= 19
238(о)= 1576	241(с)= 1576	238- 241= 3	238- 145= 93
32()= 3094	35(#)= 3094	32- 35= 3	32- 97= 65

Данные, представленные в таблице 1, наглядно демонстрируют влияние дополнительного метода кодирования по алгоритму БП на криптостойкость шифра Цезаря. В третьей колонке представлены результаты частотного анализа классического шифра Цезаря. Как видно из представленных данных, можно однозначно определить ключ шифра (в данном случае это число 3).

В четвертой колонке представлены результаты частотного анализа над шифром Цезаря+БП; в данном случае мы не можем однозначно определить ключ.

Благодаря простоте работы алгоритма комбинации простых шифров замены можно достигнуть большего быстродействия, и, следовательно, данный подход актуален для оптимизации современных программных средств шифрования.

В таблице 2 представлены результаты частотного анализа над шифром Виженера+БП.

Сводные данные сдвигов символов шифра Виженера

Эталонный текст	По алгоритму Виженера	Сдвиги символов алгоритму Виженера	Сдвиги символов по алгоритму Виженера+БП
87(W)= 4	36(\$)= 4	87- 36= 51	87- 209= 122
46(.)= 5	222(Ю)= 5	46- 222= 176	46- 131= 85
103(g)= 5	63(?)= 5	103- 63= 40	103- 33= 70
84(T)= 5	12()= 5	84- 12= 72	84- 34= 50
117(u)= 5	77(M)= 5	117- 77= 40	117- 161= 44
112(p)= 6	255(я)= 6	112- 255= 143	112- 99= 13
99(c)= 8	82(R)= 8	99- 82= 17	99- 104= 5
102(f)= 9	15()= 9	102- 15= 87	102- 31= 71
121(y)= 9	87(W)= 9	121- 87= 34	121- 90= 31
119(w)= 11	11()	119- 11= 108	119- 81= 38
108(l)= 15	101(e)= 15	108- 101= 7	108- 79= 29
100(d)= 17	191(i)= 17	100- 191= 91	100- 156= 56
115(s)= 22	167(§)= 22	115- 167= 52	115- 132= 17
104(h)= 25	51(3)= 25	104- 51= 53	104- 162= 58
114(r)= 25	241(c)= 25	114- 241= 127	114- 183= 69
105(i)= 27	61(=)= 27	105- 61= 44	105- 105= 0
111(o)= 28	6()= 28	111- 6= 105	111- 194= 83
110(n)= 31	194(B)= 31	110- 194= 84	110- 128= 18
116(t)= 37	184(ё)= 37	116- 184= 68	116- 208= 92
97(a)= 39	145(')= 39	97- 145= 48	97- 19= 78
101(e)= 48	46(.)= 48	101- 46= 55	101- 226= 125
32()= 88	182(¶)= 88	32- 182= 150	32- 67= 35

Полученный шифр не поддается простому частотному анализу. Получить исходный файл не удастся без использования более громоздких алгоритмов, а структура используемых шифров сделает этот процесс еще более трудоемким.

Использование простых шифров замены с модификатором Баттисты Порты с успехом рекомендуется использовать для шифрования конфиденциальной информации, срок актуальности которой не превышает одного месяца.

Библиографический список

1. Панасенко С.П. Алгоритмы шифрования. Спец. справочник. – СПб.: БХВ, 2009. – 576 с.
2. Никифоров С.Н. Методы защиты информации. Шифрование данных. – М.: Лань, 2018. – 160 с.
3. Сمارт Н. Криптография. – М.: Техносфера, 2005. – 528 с.

A.G. Klyuchov, V.I. Loginov, Ju.S. Fedosenko

INCREASING THE CRYPTOGRAPHIC STRENGTH OF ENCRYPTION ALGORITHMS BASED ON THE IDEA OF BATTISTA PORTA

Volga State University of Water Transport, N. Novgorod

The results of the development of new methods of encoding information for equipment for transmitting confidential data are presented.

Keywords: encryption, encoding

УДК 004.023

А.К. Халайджи

АНАЛИЗ КАЧЕСТВА РАБОТЫ ДЕТЕКТОРОВ R-ПИКОВ В СИГНАЛЕ ЭКГ В РЕЖИМЕ РЕАЛЬНОГО ВРЕМЕНИ

Московский государственный технический университет им. Н.Э. Баумана

В работе предложен алгоритм обработки сигнала ЭКГ с помощью скользящего окна, позволяющий использовать и сравнивать качество существующих детекторов R-пиков в режиме реального времени. Исследование осуществлено на сигналах MIT-BIH в соответствии с парадигмой разбиения inter-patient на обучающую и тестовую выборку. На обучающей выборке исследовано влияние медианной фильтрации на качество работы детекторов и определена наилучшая конфигурация предложенного алгоритма обработки сигнала ЭКГ в режиме реального времени. Проведён количественный и качественный анализ существующих алгоритмов обнаружения R-пиков с целью определения их достоинств и недостатков.

Ключевые слова: реальное время, скользящее окно, медианный фильтр, детектор R-пиков, MIT-BIH.

Введение

Для обеспечения возможности автоматизированного анализа сигнала ЭКГ необходимо осуществить его первичную обработку: очистить от шумов, удалить тренд и извлечь полезные характеристики сигнала для анализа. Одной из наиболее часто используемых характеристик сигнала ЭКГ являются положения R-пиков [1]. Это связано с относительной простотой их извлечения даже из зашумлённого сигнала и существованием обширного количества алгоритмов их обнаружения.

Точность анализа сильно зависит от качества процедуры первичной обработки. Во многих работах по умолчанию используется алгоритм Пана-Томпкинса [2], поскольку он работает в режиме реального времени и был исторически одним из первых наиболее удачных детекторов, однако он был представлен в 1985 году и всё же ориентирован на обнаружение QRS-комплексов, а не точных положений R-пиков. Несмотря на большое количество алгоритмов обнаружения R-пиков, не существует такого, который позволяет безошибочно обрабатывать произвольный сигнал ЭКГ, хотя бы в рамках существующих баз сигналов вида MIT-BIH. Этим обусловлена актуальность проведения дальнейших исследований в этой области с целью повышения точности обнаружения R-пиков.

Процедура объективного выбора наиболее качественных детекторов R-пиков затруднена в силу нескольких причин:

- в большинстве работ указаны только итоговые показатели точности без подробного описания схемы эксперимента;
- используются разные наборы данных для оценки качества алгоритмов.
- используются разные метрики качества для представления результатов.

В рамках исследования предложен алгоритм обработки сигнала ЭКГ в реальном времени, выбраны метрики качества, описаны алгоритмы их подсчёта и проведены эксперименты для существующих детекторов. Для анализа детекторов использовались сигналы из MIT-BIH. В результате выявлены лучшие конфигурации детекторов, выполнено их сравнение, и предложены направления дальнейших исследований.

Выбор детекторов R-пиков и метрик качества обнаружения R-пиков

Все существующие алгоритмы обладают определённой погрешностью относительно точного положения R-пика. Так, часто алгоритмы анализируют наличие сильных перепадов по значению амплитуды сигнала, которые находятся на несколько отсчётов позже или раньше непосредственного положения R-пика. Для учёта этого обстоятельства в исследовании использовался специальный параметр R , который задаёт радиус допустимой для алгоритма окрестности в отсчётах. Значение этого параметра можно определить с помощью перебора множества значений и сравнения итогового качества алгоритмов на обучающей выборке. В качестве обучающей выборки использовалось обучающее подмножество сигналов согласно парадигме разбиения *inter-patient* [3].

Для оценки качества алгоритмов использовались следующие метрики качества:

- TP – число R-пиков из разметки, в окрестности радиуса R относительно которых алгоритм обнаружил R-пик;
- FN – число R-пиков из разметки, для которых алгоритмом не было обнаружено ни одного R-пика в рамках окрестности радиуса R ;
- FP – число R-пиков, находящихся в окрестности радиуса R относительно истинных положений R-пиков, кроме ближайших к ним;
- MAE – среднее в отсчётах отклонение TP R-пиков от истинных положений;
- Acc – доля правильных ответов алгоритма (accuracy): $Acc = TP / (TP + FN + FP)$
- Se – полнота (recall): $Se = TP / (TP + FN)$
- P⁺ – точность (precision): $P^+ = TP / (TP + FP)$

- H – среднее гармоническое между Acc, Se и P⁺:
$$H = 3 \left(\frac{1}{Acc} + \frac{1}{Se} + \frac{1}{P^+} \right)^{-1}$$

Все эксперименты программно реализованы с использованием языка Python 3.8. Для исследования существующих детекторов R-пиков выбрана библиотека *py-ecg-detectors* [4], которая была составлена из наиболее часто используемых детекторов на состояние 2018 года: Pan and Tompkins [2], Hamilton [5], Christov [6], Engelse and Zeelenberg [7], Stationary Wavelet Transform [8], Two Moving Average [9]. К этому списку добавлен наивный алгоритм обнаружения R-пиков *DumbMax*, идея которого заключается в фиксации локальных максимумов сигнала, превышающих заданный порог.

Алгоритм *DumbMax* требует задания двух параметров: R – минимальное допустимое количество отсчётов между соседними R-пиками, и порог – минимальное значение амплитуды по модулю, необходимое для фиксации отсчёта в качестве R-пика. Для корректной работы алгоритма необходимо предварительно удалить тренд из сигнала. Алгоритм последовательно рассматривает отсчёт за отсчётом, обновляя текущее значение локального максимума значения амплитуды сигнала по модулю. Если спустя R отсчётов значение не изменяется, то оно сравнивается с порогом – при превышении соответствующий ему отсчёт считается зафиксированным R-пиком, иначе – пик пропускается, и начинается новый поиск локального максимума. Сравнение с порогом используется для отсеки шума и предотвращения ложных срабатываний алгоритма в связи с ограниченностью значения окна.

Описание механизма использования детекторов в реальном времени

Записи ЭКГ в МГТ-ВН в среднем имеют длительность 30 минут и представлены в виде дискретной последовательности отсчётов на частоте 360 Гц. Каждый из детекторов требует на входе фрагмент сигнала для анализа. Если длительность этого сигнала слишком мала (например, будет содержать всего 1 R-пик), то не все детекторы смогут его корректно распознать. С другой стороны, размер анализируемого интервала не должен быть слишком большим, т.к. предпочтительно выполнять анализ в режиме реального

времени. Поэтому для дальнейших экспериментов размер временного окна, на котором запускались детекторы, выбран равным 5 секундам (или 1800 отсчётам).

Поскольку окно статической длительности не учитывает морфологию сигнала, то существуют ошибки определения крайних R-пиков – первого и последнего. Для решения этого недостатка следует перемещать окно с перекрытием предыдущего положения как минимум на один R-пик, который уже зафиксирован, чтобы минимизировать вероятность пропуска положения следующего за ним R-пика. С другой стороны, в рамках одного окна не имеет смысла перемещать окно после каждого найденного R-пика, т.к. это значительно повышает вычислительную сложность за счёт повторного перевычисления положений R-пиков каждым из детекторов, начиная со следующего после обнаруженного. Поэтому выбрана следующая схема перемещения окна для каждого из детекторов:

- Изначально окно устанавливается на нулевой отсчёт сигнала.
- Для окна запускается детектор R-пиков, возвращающий список отсчётов.
- Если хотя бы один R-пик уже был зафиксирован, то из списка удаляются все те отсчёты, которые находятся раньше отсчёта, задаваемого положением последнего зафиксированного R-пика, смещённого вправо на значение радиуса R .
- Окно сдвигается, после чего процедура повторяется в шаге 2. Новое положение начала окна определяется максимальным значением из двух возможных:
 - текущее положение, сдвинутое вправо на R ;
 - положение последнего зафиксированного R-пика за вычетом R .

Некоторые детекторы (например, DumbMax) требуют предварительной очистки сигнала от шумов и удаления тренда. Для выполнения обеих этих процедур использовалась схема из двух медианных фильтров, предложенная в [10]. Выбор обусловлен тем, что результаты исследования использовались в дальнейшем для задачи автоматической классификации аритмий, и согласно [3, 10, 11] рекомендуется использовать именно эту схему предобработки сигналов для возможности сравнения результатов с другими работами.

Для фильтрации в режиме реального времени (т.е. по мере поступления новых отсчётов), использовался временный буфер, в котором использовались значения из фрагмента сигнала в скользящем окне с добавленными к нему слева и справа значениями исходного необработанного сигнала, необходимыми для корректной работы медианных фильтров. Таким образом, алгоритму обработки сигнала доступно сразу два сигнала: исходный и после обработки медианными фильтрами. Это позволило провести анализ влияния фильтрации на качество работы детекторов.

Выбор наилучшей конфигурации детекторов на обучающей выборке

Для сравнения детекторов необходимо выбрать единое значение параметра R – минимального допустимого количества отсчётов между соседними положениями R-пиков. Понятнее всего интерпретировать значение этого параметра можно на примере наивного алгоритма DumbMax. При малых значениях R алгоритм потенциально может обнаружить больше R-пиков, в том числе и ложно, а слишком большие значения приведут к пропускам.

Для разных детекторов лучшее значение параметра R может быть разным – например, если алгоритм склонен к ложным срабатываниям, то при малых значениях R его качество будет хуже. Тем не менее, для возможности сравнения детекторов необходимо выбрать единое значение этого параметра, т.к. все они анализируют один и тот же сигнал.

В [2] приведено минимальное значение для нормальной физиологической длительности между соседними QRS-комплексами, равное 200 мс, а также средней длительности одного QRS-комплекса, равное 150 мс. С учётом этой информации, анализ качества работы детекторов при разных значениях параметра R в диапазоне от 50 до 150 отсчётов, что соответствует при частоте 360 Гц диапазону от 138 мс до 417 мс. Также

исследовалось влияние фильтрации на качество работы детекторов. Наконец, для DumbMax-детектора были проанализированы четыре разных значения порога: 50, 100, 150 и 200. Нулевое значение порога не исследовалось, т.к. оно априори создавало бы много ложных срабатываний, поскольку фиксировало бы R-пик практически через каждые R отсчётов. В табл. 1 представлены значения метрик качества для лучших конфигураций детекторов, а на рис. 1 – график гармонического среднего от значения параметра R .

Таблица 1

Лучшие конфигурации детекторов на обучающей выборке

Конфигурация	TP	FP	FN	MAE	Acc, %	Se, %	P ⁺ , %	H, %
Christov, с филтвр., $R=80$	50980	89	505	3.30	98.85	99.02	99.83	99.23
Two Moving Average, $R=95$	51027	303	458	24.50	98.53	99.11	99.41	99.02
Pan and Tompkins, $R=115$	51057	366	428	20.67	98.47	99.17	99.29	98.98
DumbMax, порог=100 с филтвр., $R=115$	50665	82	820	1.67	98.25	98.41	99.84	98.83
Stationary Wavelet Transform $R=105$	50998	587	487	13.40	97.94	99.05	98.86	98.61
Hamilton, $R=105$	50724	622	761	19.11	97.35	98.52	98.79	98.22
DumbMax, порог=50 с филтвр., $R=140$	50775	731	710	4.19	97.24	98.62	98.58	98.14
DumbMax, порог=150 с филтвр., $R=110$	48832	41	2653	1.29	94.77	94.85	99.92	96.45
Engelse and Zeelenberg, $R=60$	44706	174	6779	1.51	86.54	86.83	99.61	90.60
DumbMax, порог=200 с филтвр., $R=105$	43015	18	8470	1.20	83.52	83.55	99.96	88.38

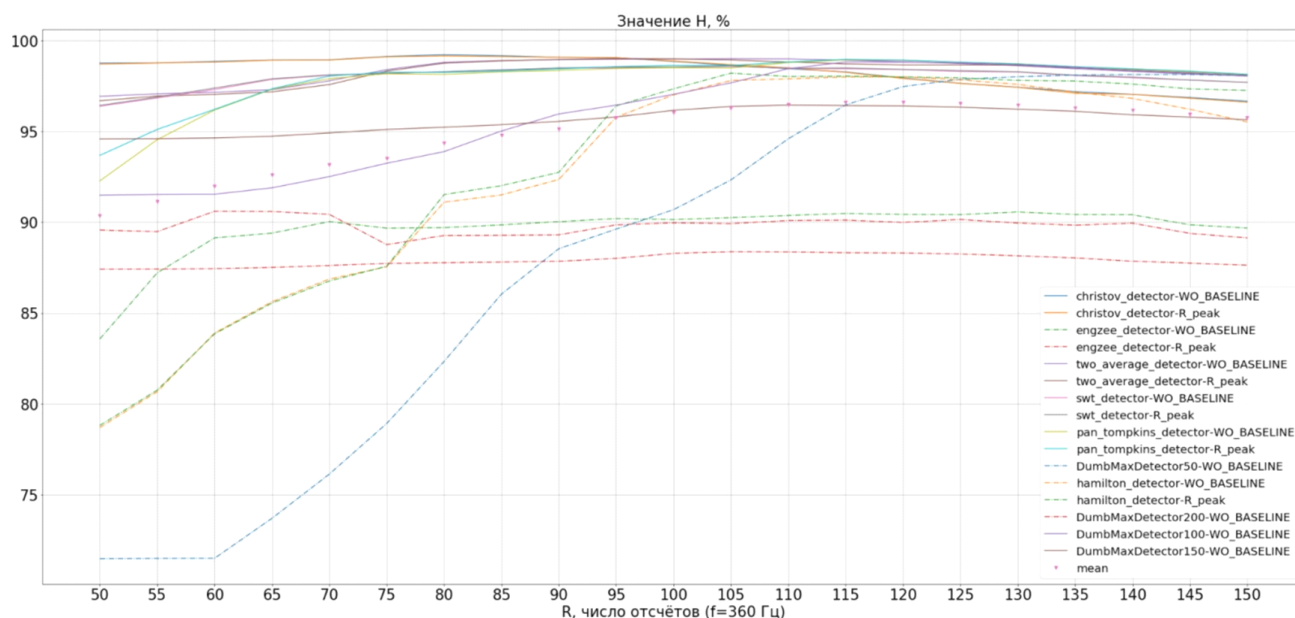


Рис. 1. Зависимость гармонического среднего для конфигураций детекторов от параметра R

Из рис. 1 следует, что лучшим «в среднем» для всех детекторов является значение R , равное 115 отсчётам или 320 мс, что согласуется с приведёнными физиологическими параметрами в [2]. Несмотря на то, что качество некоторых детекторов ниже для этого значения параметра R (например, по сравнению с Christov для R , равного 80), использование такого значения более целесообразно, т.к. больше алгоритмов работает качественно.

Также из рис. 1 видно, что алгоритмы DumbMax с порогом 200 и Engelse and Zeelenberg работают намного хуже, чем все остальные детекторы, допуская очень много пропусков (FN). Это обусловлено малыми амплитудами исследуемых сигналов и искажённостью из-за наличия аритмий и шумов, что оказалось критическим для алгоритма Engelse and Zeelenberg. Поэтому оба алгоритма исключены из дальнейшего анализа. Наконец, из рис. 1 следует, что фильтрация практически не влияет на остальные алгоритмы. Поэтому целесообразно использовать только одну конфигурацию из них, которая при определённом значении параметра R позволяет достичь лучшего качества.

Анализ итоговых конфигураций детекторов

В табл. 2–4 представлены результаты работы отобранных конфигураций детекторов на обучающей, тестовой и обеих выборках соответственно.

Таблица 2

Лучшие конфигурации детекторов на обучающей выборке для $R=115$

Конфигурация	TP	FP	FN	MAE	Acc, %	Se, %	P+, %	H, %
Christov с фил. $R=80$	50980	89	505	3.30	98.85	99.02	99.83	99.23
Pan and Tompkins	51057	366	428	20.67	98.47	99.17	99.29	98.98
Two Moving Average с фил	50956	295	529	24.55	98.41	98.97	99.42	98.93
DumbMax порог=100 с фил	50665	82	820	1.67	98.25	98.41	99.84	98.83
Stationary Wavelet Transform	50909	585	576	13.63	97.77	98.88	98.86	98.50
Christov с фил	50195	23	1290	4.68	97.45	97.49	99.95	98.28
Hamilton	50163	172	1322	19.59	97.11	97.43	99.66	98.05
DumbMax порог=50 с фил	51155	2483	330	3.57	94.79	99.36	95.37	96.46
DumbMax порог=150 с фил	48794	18	2691	1.32	94.74	94.77	99.96	96.43

Таблица 3

Лучшие конфигурации детекторов на тестовой выборке для $R=115$

Конфигурация	TP	FP	FN	MAE	Acc, %	Se, %	P+, %	H, %
Hamilton	49396	179	309	18.46	99.02	99.38	99.64	99.35
Pan and Tompkins	49571	413	134	18.58	98.91	99.73	99.17	99.27
Christov с фил. $R=80$	49048	137	657	2.81	98.41	98.68	99.72	98.93
Stationary Wavelet Transform	49495	740	210	11.46	98.12	99.58	98.53	98.74
Two Moving Average с фил	49593	1204	112	23.86	97.41	99.77	97.63	98.26
Christov с фил	48380	83	1325	3.35	97.17	97.33	99.83	98.10
DumbMax порог=100 с фил	49089	1576	616	1.11	95.73	98.76	96.89	97.11
DumbMax порог=50 с фил	49661	4121	44	1.60	92.26	99.91	92.34	94.70
DumbMax порог=150 с фил	45466	1161	4239	1.03	89.38	91.47	97.51	92.66

Таблица 4

Лучшие конфигурации детекторов на всех сигналах для R=115

Конфигурация	TP	FP	FN	MAE	Acc, %	Se, %	P+, %	H, %
Pan and Tompkins	101125	778	562	19.48	98.69	99.45	99.24	99.13
Christov с фил. R=80	100526	226	1161	3.02	98.64	98.86	99.78	99.09
Hamilton	100054	350	1633	18.98	98.06	98.39	99.65	98.70
Stationary Wavelet Transform	100899	1323	788	12.44	97.95	99.23	98.71	98.63
Two Moving Average с фил	101044	1499	643	24.25	97.92	99.37	98.54	98.61
Christov с фил	99070	106	2617	3.92	97.32	97.43	99.89	98.20
DumbMax порог=100 с фил	100251	1654	1436	1.37	97.01	98.59	98.38	97.99
DumbMax порог=50 с фил	101312	6583	375	2.55	93.57	99.63	93.90	95.62
DumbMax порог=150 с фил	94756	1178	6931	1.15	92.12	93.18	98.77	94.60

Полученные результаты показывают, что действительно алгоритм Пана-Томпкинса действительно является одним из лучших алгоритмов для сигналов из MIT-BIH, но обладает большим значением MAE, т.к. находит QRS-комплексы, а не точные положения R-пиков. Наивные алгоритмы DumbMax обладают большим количеством ложных срабатываний, однако достигают наилучшего значения MAE, а при пороге 50 – и FN. Аналогичными свойствами обладает алгоритм Christov. Из таблиц видно, что при меньшем значении параметра R алгоритм достигает лучших показателей качества в силу меньшего количества пропусков, однако при этом допуская больше ложных срабатываний. Алгоритм Two Moving Average обладает довольно высокой полнотой, однако допускает много ложных срабатываний на сигналах из тестовой выборки и обладает большим значением MAE. Достоинствами алгоритма Stationary Wavelet Transform можно выделить его скорость работы и относительно малое количество пропусков, однако он сильно чувствителен к нарушениям морфологии аритмиями, из-за чего допускает большое количество ложных срабатываний. Наконец, алгоритм Hamilton обладает одним из лучших показателей по числу ложных срабатываний, однако имеет большую погрешность и пропускает большое количество пиков.

Заключение

Подтверждена эффективность алгоритма Пана-Томпкинса для сигналов из MIT-BIH, однако выявлено, что этот алгоритм неточно определяет положение R-пика внутри QRS-комплекса и допускает много ложных срабатываний. В свою очередь, другие детекторы устраняют часть из этих недостатков, например, Christov практически не допускает ложных срабатываний, а DumbMax-детекторы наиболее точно определяют положение R-пиков. При этом не существует единого алгоритма, который как на обучающей выборке, так и на тестовой и всех сигналах сразу показал бы явный отрыв по качеству от остальных. Наконец, каждый из алгоритмов хуже остальных согласно хотя бы одной из метрик качества.

Полученные результаты приводят к идее совместного использования результатов сразу нескольких детекторов для устранения недостатков и улучшения сильных сторон друг друга. Другой актуальной задачей является разработка алгоритмов для уточнения итоговых положений R-пиков, что важно для извлечения морфологических признаков R-пиков и определения длительностей RR-интервалов для автоматизированного анализа ЭКГ.

Библиографический список

1. Кузнецов А.А. Биофизика сердца. Методы обработки и анализа электрокардиографической информации при донозологических исследованиях: учеб. пособие / А.А. Кузнецов ; Владим. гос. ун-т имени Александра Григорьевича и Николая Григорьевича Столетовых. – Владимир: Изд-во ВлГУ, 2012. – 237 с.
2. Pan J. A Real-Time QRS Detection Algorithm / J. Pan, W.J. Tompkins // IEEE Transactions on Biomedical Engineering BME-32.3, 1985. – P. 230–236. DOI: 10.1109/TBME.1985.325532
3. Luz E.J.S. ECG-based heartbeat classification for arrhythmia detection: A survey / E.J.S. Luz, W.R. Schwartz, G. Camara-Chavez, D. Menotti // Computer Methods and Programs in Biomedicine, 2016. – V. 127 – P. 144-164. DOI: 10.1016/j.cmpb.2015.12.008
4. Porr B. Popular ECG R peak detectors written in python (Version 1.0.2) / B. Porr // Zenodo, 2019. DOI: 10.5281/zenodo.3588108
5. Hamilton P.S. Open Source ECG Analysis Software Documentation / P.S. Hamilton // Computers in Cardiology, 2002. – V. 29. – P. 101-104. DOI: 10.1109/CIC.2002.1166717
6. Christov I.I. Real time electrocardiogram QRS detection using combined adaptive threshold / I.I. Christov // BioMedical Engineering OnLine, 2004. – V. 3. – №28. – P. 9. DOI: 10.1186/1475-925X-3-28
7. Lourenco A. Real Time Electrocardiogram Segmentation for Finger Based ECG Biometrics / A. Lourenco, H. Silva, P. Leite, R. Lourenco, A. Fred // BIOSIGNALS, 2012. – P. 49-54. DOI: 10.5220/0003777300490054
8. Kalidas V. Real-time QRS detector using Stationary Wavelet Transform for Automated ECG Analysis / V. Kalidas, L. Tamil // 2017 IEEE 17th International Conference on Bioinformatics and Bioengineering (BIBE), 2017. – P. 457-461. DOI: 10.1109/BIBE.2017.00-12
9. Elgendi M. Frequency Bands Effects on QRS Detection / M. Elgendi, M. Jonkman, F. De Boer // The 3rd International Conference on Bio-inspired Systems and Signal Processing (BIOSIGNALS2010), 2010. – V. 1. – P. 428-431. DOI: 10.5220/0002742704280431
10. de Chazal P. Automatic classification of heartbeats using ECG morphology and heartbeat interval features / P. de Chazal, M. O'Dwyer, R.B. Reilly // IEEE Trans. Biomed. Eng. – 2004. – V. 51. – № 7. – P. 1196–1206. DOI: 10.1109/tbme.2004.827359
11. Аль-Салех Г.И. Сравнительный анализ методов обработки электрокардиографического сигнала / Г.И. Аль-Салех, А.А. Рунова, А.В. Писарева, А.П. Николаев // Физика и радиоэлектроника в медицине и экологии. Труды XIII Международной научной конференции с научной молодёжной школой имени И.Н. Спиридонова. В 2-х кН., 2018, Владимир-Суздаль. – Изд-во ООО «Графика». – С.120-125.

A.K. Khalaydzh

ANALYSIS OF THE QUALITY OF R-PEAK DETECTORS IN ECG SIGNAL IN THE REAL-TIME MODE

Bauman Moscow state technical university, Moscow

Purpose: The author provides an algorithm of ECG signal processing via sliding window to use and compare existing R-peak detectors in real-time mode.

Design / methodology / approach: The paper chooses quality metrics to compare modern R-peak detectors from the Python module py-ecg-detectors. It also describes a dumb max detector that consequently searches local maximal value and trims it via threshold. Author introduces important parameter of radius R , which can be interpreted as a minimal distance between adjacent R-peaks or an appropriate radius of the area, surrounding truth positions of R-peaks, to take into account the tolerance of R-peak detectors. Paper also describes an algorithm of scanning and filtering signal in real-time mode via sliding window and explains how to decrease errors when moving it. Author uses inter-patient paradigm of MIT-BIH signal division to find out the best detectors configurations on train subset and evaluate final quality on test and all signals.

Findings: The paper finds out optimal value of optimal radius R in sense of mean for the majority of R-peak detectors. It also makes conclusions about advantages and disadvantages of each of detectors based on the chosen quality metrics.

Research limitations/implications: Research is limited to MIT-BIH signals database and using detectors only from Python module py-ecg-detectors. Nevertheless, the obtained results can be further used in systems of automatic ECG analysis such as arrhythmia detection.

Originality/value: The value of the whole paper is the uniform algorithm and environment that was used to compare existing R-peak detectors in real-time mode. In conclusion the paper shows that there is no single detector from analyzed set of detectors in this research that is the best on train, test and overall signals subsets simultaneously, based on chosen quality metrics. The paper also notes that for each R-peak detector there is at least one quality metric that is worse than the others obtain. As a conclusion the author suggests two perspective research directions to increase the final quality of the analyzed procedure or R-peak detection. First is joint using the results, obtained by several detectors simultaneously, to decrease their disadvantages and increase their advantages. The second one is about the development of new algorithms, adjusting the final R-peak positions, found by R-peak detectors, due to its importance to the problems of extraction of the R-peak morphological features or RR-intervals durations.

Keywords: real-time mode, sliding window, median filter, R-peak detector, MIT-BIH

АЛГОРИТМ ОБНАРУЖЕНИЯ R-ПИКОВ СИГНАЛА ЭКГ НА ОСНОВЕ КЛАСТЕРИЗАЦИИ РЕЗУЛЬТАТОВ РАБОТЫ ОТДЕЛЬНЫХ ДЕТЕКТОРОВ

Московский государственный технический университет им. Н.Э. Баумана

В статье предложен алгоритм одномерной адаптивной кластеризации, использующий реализацию KMeans на основе динамического программирования. На его основе разработан алгоритм объединения результатов работы отдельных детекторов R-пиков в режиме реального времени. Осуществлён отбор лучших конфигураций объединения детекторов на обучающем множестве сигналов согласно парадигме разбиения inter-patient сигналов из MIT-BIH. Показано, что использование нескольких детекторов действительно повышает качество обнаружения R-пиков. Выявлена конфигурация из четырёх детекторов, которая обладает лучшим качеством по сравнению с остальными как на обучающей выборке, так и на тестовой и всех сигналах сразу.

Ключевые слова: одномерная адаптивная кластеризация, KMeans, детектор R-пиков, MIT-BIH.

Введение

Точность автоматизированного анализа ЭКГ сильно зависит от качества процедуры первичной обработки. Одной из наиболее часто используемых характеристик сигнала являются положения R-пиков [1]. В рамках предварительного исследования [2] был создан алгоритм обработки сигнала ЭКГ, использующий скользящее окно длительностью 5 секунд, для возможности использования и сравнения детекторов R-пиков в режиме реального времени. Был проведён анализ реализаций часто используемых детекторов R-пиков на состояние 2018 года, собранных в Python-модуль py-ecg-detectors. [3]: Pan and Tompkins [4], Hamilton [5], Christov [6], Engelse and Zeelenberg [7], Stationary Wavelet Transform [8], Two Moving Average [9]. Также был предложен наивный алгоритм DumbMax, отслеживающий локальные максимумы и фиксирующий их при превышении порога. Для оценки качества введён параметр R – радиус допустимой окрестности вокруг истинного положения R-пика, в рамках которой пик считается распознанным алгоритмом верно. Использовались следующие метрики качества:

- TP – число R-пиков из разметки, в окрестности радиуса R относительно которых алгоритм обнаружил R-пик;
- FN – число R-пиков из разметки, для которых алгоритмом не было обнаружено ни одного R-пика в рамках окрестности радиуса R ;
- FP – число R-пиков, находящихся в окрестности радиуса R относительно истинных положений R-пиков, кроме ближайших к ним;
- MAE – среднее в отсчётах отклонение TP R-пиков от истинных положений;
- Acc – доля правильных ответов алгоритма (accuracy):

$$Acc = TP / (TP + FN + FP)$$

- Se – полнота (recall): $Se = TP / (TP + FN)$

- P^+ – точность (precision): $P^+ = TP / (TP + FP)$

- H – среднее гармоническое между Acc, Se и P^+ :

$$H = 3 \left(\frac{1}{Acc} + \frac{1}{Se} + \frac{1}{P^+} \right)^{-1}$$

В [2] использовалось разделение на обучающую и тестовую выборку согласно парадигме разбиения inter-patient сигналов из MIT-BIH [10]. Показано, что лучшим значением параметра R является 115 отсчётов. Полученные в [2] результаты для

обучающей выборки представлены в табл. 1. Для некоторых детекторов указана необходимость предварительной фильтрации сигнала (в [2] использовались медианные фильтры, как в [11], эффективность которых подтверждена и в [12]).

Таблица 1

Лучшие конфигурации детекторов на обучающей выборке

№	Конфигурация	TP	FP	FN	MAE	Acc, %	Se, %	P+, %	H, %
1	Pan and Tompkins	51057	366	428	20.67	98.47	99.17	99.29	98.98
2	Two Moving Average с фил	50956	295	529	24.55	98.41	98.97	99.42	98.93
3	DumbMax порог=100 с фил	50665	82	820	1.67	98.25	98.41	99.84	98.83
4	Stationary Wavelet Transform	50909	585	576	13.63	97.77	98.88	98.86	98.50
5	Christov с фил	50195	23	1290	4.68	97.45	97.49	99.95	98.28
6	Hamilton	50163	172	1322	19.59	97.11	97.43	99.66	98.05
7	DumbMax порог=50 с фил	51155	2483	330	3.57	94.79	99.36	95.37	96.46
8	DumbMax порог=150 с фил	48794	18	2691	1.32	94.74	94.77	99.96	96.43

Настоящая статья нацелена решить те недостатки, которые были приведены в заключении [2]. Для этого представлен оригинальный подход объединения результатов работы существующих детекторов с помощью алгоритма адаптивной одномерной кластеризации KMeans. Предложенный подход позволил снизить количество ложно обнаруживаемых R-пиков и повысить итоговое качество обнаружения положений R-пиков.

Алгоритм совместного обнаружения R-пиков с учётом информации от нескольких детекторов

Для уменьшения числа ложных срабатываний используется следующая идея: если больше половины детекторов обнаружили один и тот же R-пик, то он фиксируется, иначе – игнорируется. Основная сложность заключается в фиксации детекторами разных отсчётов для одних и тех же R-пиков. Для её преодоления необходимо объединять друг с другом близкие отсчёты R-пиков, обнаруженные разными детекторами, а итоговой зафиксированной позицией R-пика считать их среднее значение (при этой операции может получиться нецелый отсчёт, но это решается округлением до ближайшего целого).

Таким образом, для объединения результатов необходимо решить следующую задачу:

- имеется список упорядоченных по отсчётам массивов положений R-пиков;
- требуется сформировать итоговый массив положений R-пиков, объединив близкие R-пики из разных массивов в группы, и оставив только те из них, в которых содержатся отсчёты более чем из половины массивов исходного списка;
- каждую такую группу близких значений – заменить средним по группе;
- расстояние между итоговыми соседними R-пиками должно быть не меньше R .

Формулировка напоминает постановку задачи сортировки слиянием, в которой вместо точного сравнения допускается сравнение с отклонением в R . Однако такой подход к решению задачи будет реализовывать жадную стратегию, которая, к сожалению, не будет достигать оптимального решения. Это можно продемонстрировать на следующем примере. Пусть $R = 75$, и на вход алгоритму поступил следующий список для двух детекторов: $[[0, 75], [0, 76]]$. Жадный алгоритм слияния объединит положения 0 и 75, т.к.

они находятся в радиусе R , однако отсчёт 76 уже к ним не отнесёт, что приведёт к итоговому массиву [(25, 3), (76, 1)], т.к. первое значение было получено усреднением трёх точек, а второе – одной. При этом ожидаемым результатом являлся другой массив [(0, 2), (75.5, 2)]. Таким образом, жадная стратегия не позволяет достичь оптимального результата.

Решить недостаток жадной стратегии позволяет алгоритм кластеризации. Исходная задача является одномерной, если все положения из разных массивов нанести на числовую ось, например, с помощью базового алгоритма сортировки слиянием. Тогда объединение результатов разных детекторов можно получить, кластеризовав все отсчёты на числовой оси. В таком случае каждый кластер характеризуется количеством отсчёты, вошедших в него, и центром. Если кластер состоит только из отсчётов одного детектора, смысла в таком кластере нет, однако если в нём присутствуют отсчёты из более чем половины детекторов, то центр такого кластера может быть интерпретирован как уточнённое положение R -пика.

Для реализации подхода необходимо, чтобы алгоритм кластеризации достигал глобально оптимальное решение, не требовал явного задания априорного количества кластеров и обладал гарантиями сходимости. Классический алгоритм KMeans не соответствует ни одному из этих требований. Однако существует реализация алгоритма KMeans1D [13], которая использует механизм динамического программирования и обладает первым и последним свойствами для одномерных задач – она позволяет оптимально кластеризовать n точек на k кластеров с вычислительной сложностью $O(kn^2)$.

От необходимости явного задания априорного количества кластеров можно избавиться с помощью добавления в алгоритм механизма ранней остановки при выполнении определённых условий, проверяемых на каждом шаге. Тогда, адаптивная кластеризация заключается в последовательном увеличении числа кластеров от 1 до n , где n – изначальное число точек, или до выполнения условия остановки. В отличие от классического алгоритма KMeans, структура динамического программирования позволяет эффективно осуществлять перерасчёт для увеличенного на один числа кластеров, используя текущие результаты.

В рамках анализируемого окна в 5 секунд количество R -пиков можно оценить в 5–15 (для пульса от 60 до 180 уд/мин). Поскольку используется не более 10 детекторов, то количество всех кластеризуемых отсчётов можно оценить числом 150. Итоговое количество кластеров, в среднем, совпадает с истинным количеством R -пиков. Поэтому для каждых 5 секунд сигнала необходимо в худшем случае выполнять не более чем порядка $kn^2 \propto 15 \cdot 150^2 = 337500$ операций, что может быть выполнено быстрее, чем за секунду на языке Python. Поэтому алгоритм может выполняться в режиме реального времени.

Для возможности использования алгоритма адаптивной кластеризации необходимо определить эвристику ранней остановки. К кластерам предъявляются 2 требования:

- если есть два кластера, расстояние между центрами, которых больше R , то кластеризация завершается;
- внутри одного кластера, содержащего более 1 отсчёта, не должно быть точек только от одного детектора. При обнаружении такой ситуации, кластеризация продолжается с увеличенным числом кластеров, т.к. необходимо продолжить разбиение кластера;

Первое правило выполняется для одного кластера, в котором сосредоточены все исходные точки. Если при увеличении числа кластеров от k до $k+1$ первое правило нарушилось, то последней подходящей кластеризацией является та, которая была получена для k кластеров. Если правило так и не нарушилось для n кластеров, где n – число точек, то результатом является исходное множество точек, каждая из которых задаёт свой кластер.

Таким образом, итоговый алгоритм состоит из следующих шагов:

- для текущего окна запускается K детекторов. В итоге получается список из K массивов с N_1, N_2, \dots, N_K отсчётами – положениями R-пиков, согласно детекторам;
- с помощью алгоритма сортировки слиянием все отсчёты наносятся на единую числовую ось, формируя в итоге упорядоченный массив из $\sum_{i=1}^K N_i$ отсчётов;
- для полученного массива запускается процедура адаптивной одномерной кластеризации от 1 до $\sum_{i=1}^K N_i$ кластеров с учётом описанной эвристики ранней остановки;
- по итогам предыдущего шага формируется список центров кластеров и их размеров, который обрабатывается последовательно следующим образом: если центр текущего кластера находится на расстоянии большем чем R (при условии наличия ранее зафиксированных R-пиков), при этом размер кластера больше $K/2$, то центр кластера фиксируется алгоритмом как положение очередного R-пика;
- после завершения анализа списка скользящее окно сдвигается в соответствии со схемой, предложенной в [2].

Исследование качества предложенного алгоритма

В табл. 2 представлены результаты на обучающей выборке для лучших 10 конфигураций из нескольких детекторов (первые 10 строк).

Таблица 2

Общие результаты для обучающей выборки

Детекторы (из табл. 1)								TP	FP	FN	MAE	Acc,%	Se,%	P+,%	H,%
1	2	3	4	5	6	7	8								
+	+			+	+	+		51177	98	308	19.04	99.21	99.40	99.81	99.473
+	+			+		+		51142	62	343	18.98	99.21	99.33	99.88	99.472
+	+		+	+	+	+		51168	102	317	18.14	99.19	99.38	99.80	99.456
+	+	+						51145	82	340	16.74	99.18	99.34	99.84	99.453
+	+		+	+	+	+	+	51160	103	325	16.14	99.17	99.37	99.80	99.446
+	+	+	+	+	+	+	+	51095	68	390	15.16	99.11	99.24	99.87	99.406
	+					+		51098	89	387	25.48	99.08	99.25	99.83	99.386
+		+	+	+	+	+	+	51072	106	413	14.49	98.99	99.20	99.79	99.326
	+	+	+	+	+		+	50953	34	532	11.53	98.90	98.97	99.93	99.264
+						+		51111	207	374	22.97	98.88	99.27	99.60	99.249
+	+							50944	67	541	21.38	98.82	98.95	99.87	99.211
			+			+		51069	274	416	16.61	98.67	99.19	99.47	99.109
		+		+	+	+	+	50997	227	488	13.26	98.62	99.05	99.56	99.075
					+	+		51016	276	469	22.90	98.56	99.09	99.46	99.035
+		+						50762	57	723	13.64	98.49	98.60	99.89	98.990
+								51057	366	428	20.67	98.47	99.17	99.29	98.975
	+		+					50742	72	743	18.22	98.42	98.56	99.86	98.942
	+							50956	295	529	24.55	98.41	98.97	99.42	98.932
+			+					50911	254	574	16.76	98.40	98.89	99.50	98.928
				+		+		50754	116	731	15.42	98.36	98.58	99.77	98.899
		+						50665	82	820	1.67	98.25	98.41	99.84	98.828
		+	+					50642	66	843	9.66	98.24	98.36	99.87	98.818
	+	+						50588	22	897	15.79	98.22	98.26	99.96	98.807
		+		+		+	+	50723	326	762	11.92	97.90	98.52	99.36	98.590
		+			+			50412	41	1073	13.00	97.84	97.92	99.92	98.551

			+					50909	585	576	13.63	97.77	98.88	98.86	98.501
		+				+	+	50733	528	752	13.57	97.54	98.54	98.97	98.346
				+				50195	23	1290	4.68	97.45	97.49	99.95	98.283
		+				+		50746	616	739	17.07	97.40	98.56	98.80	98.250
+						+		50194	155	1291	19.92	97.20	97.49	99.69	98.114
	+			+				50036	21	1449	13.48	97.15	97.19	99.96	98.083
						+		50163	172	1322	19.59	97.11	97.43	99.66	98.054
			+		+			50150	166	1335	16.08	97.09	97.41	99.67	98.043
+				+				50031	45	1454	11.61	97.09	97.18	99.91	98.043
	+				+			50030	44	1455	21.41	97.09	97.17	99.91	98.039
		+		+				49934	20	1551	5.84	96.95	96.99	99.96	97.947

Окончание табл. 2

			+	+				49677	48	1808	8.42	96.40	96.49	99.90	97.570
				+	+			49261	46	2224	10.85	95.59	95.68	99.91	97.019
+							+	48980	17	2505	11.26	95.10	95.13	99.97	96.680
		+					+	49004	133	2481	5.02	94.94	95.18	99.73	96.567
							+	51155	2483	330	3.57	94.79	99.36	95.37	96.464
							+	48794	18	2691	1.32	94.74	94.77	99.96	96.429
	+						+	48696	10	2789	13.59	94.56	94.58	99.98	96.307
						+	+	49208	571	2277	16.58	94.53	95.58	98.85	96.285
			+				+	48675	24	2810	7.66	94.50	94.54	99.95	96.263
					+		+	48147	21	3338	10.64	93.48	93.52	99.96	95.559
				+			+	47947	9	3538	3.52	93.11	93.13	99.98	95.300

Для лучших (отмечены зелёным цветом) и худших (отмечены красным цветом) конфигураций из $k=2...8$ детекторов и для отдельных детекторов из табл. 1 (отмечены серым цветом).

Анализ полученных значений из табл. 2 приводит к следующим выводам:

- конфигурации, использующие несколько детекторов, лучше. Наилучшие показатели были достигнуты для 5 и 4 детекторов; увеличение числа детекторов не всегда приводит к улучшению итоговой точности; в лучшие детекторы чаще всего входили детекторы Pan and Tompkins, Two Moving Average с фил., и DumbMax с фил., и порогом 50;

- во многих случаях объединение детекторов приводит к уменьшению числа ложных срабатываний, что снижает Se и повышает P^+ , что соответствует исходной идее;

- после объединения результатов отдельных детекторов с большим значением метрики MAE (Pan and Tompkins или Two Moving Average), этот показатель уменьшается, однако всё же остаётся довольно большим;

- Худшей конфигурацией из трёх моделей является объединение DumbMax-детекторов. Видно, что качество такой конфигурации лучше, чем у отдельных DumbMax-детекторов с порогами 150 и 50 – в ней намного меньше FP, чем в DumbMax с порогом 50, и намного меньше FN, чем в DumbMax с порогом 150. Худшее качество было у конфигураций, использующих пары детекторов, одним из которых являлся DumbMax с порогом 150 – из-за его большого количества FN, что приводило к тому, что для большого числа R-пигов не набиралось необходимое большинство для фиксации R-пика. В частности, из-за этого объединение из трёх DumbMax-детекторов хуже по качеству, чем DumbMax с порогом 100. Это подчёркивает, что при использовании описанного подхода очень важно, чтобы все модели, входящие в объединение, были примерно одного уровня по качеству, т.к. в противном случае качество объединения становится хуже, чем у лучших его компонент.

В табл. 3 и 4 приведены итоговые результаты для тестовой выборки и всех анализируемых сигналов. Полученные результаты подтверждают основные выводы для обучающей выборки. Лучшим детектором для одиночных детекторов оказался Pan and Tompkins, а лучшей конфигурацией из нескольких детекторов – объединение Pan and Tompkins, Two Moving Average с фил., Christov с фил. и DumbMax с фил. и порогом 50.

Таблица 3

Результаты для лучших конфигураций на тестовой выборке

Детекторы (из табл. 1)								TP	FP	FN	MAE	Acc,%	Se,%	P+,%	H,%
1	2	3	4	5	6	7	8								
+	+			+		+		49629	175	76	18.88	99.50	99.85	99.65	99.666
+	+	+	+	+	+	+	+	49614	164	91	15.90	99.49	99.82	99.67	99.660
	+					+		49643	217	62	26.99	99.44	99.88	99.56	99.626
+	+		+	+	+	+		49629	212	76	17.35	99.42	99.85	99.57	99.613
+	+		+	+	+	+	+	49631	218	74	16.01	99.42	99.85	99.56	99.610
+	+			+	+	+		49644	240	61	18.60	99.40	99.88	99.52	99.600
+	+	+						49638	306	67	17.78	99.25	99.87	99.39	99.503

Окончание табл. 3

					+			49396	179	309	18.46	99.02	99.38	99.64	99.346
+								49571	413	134	18.58	98.91	99.73	99.17	99.269
			+					49495	740	210	11.46	98.12	99.58	98.53	98.740
	+							49593	1204	112	23.86	97.41	99.77	97.63	98.259
				+				48380	83	1325	3.35	97.17	97.33	99.83	98.095
		+						49089	1576	616	1.11	95.73	98.76	96.89	97.111
						+		49661	4121	44	1.60	92.26	99.91	92.34	94.704
							+	45466	1161	4239	1.03	89.38	91.47	97.51	92.661

Таблица 4

Результаты для лучших конфигураций на всех сигналах

Детекторы (из табл. 1)								TP	FP	FN	MAE	Acc,%	Se,%	P+,%	H,%
1	2	3	4	5	6	7	8								
+	+			+		+		101266	237	421	18.67	99.35	99.59	99.77	99.570
+	+			+	+	+		101316	339	371	18.60	99.30	99.64	99.67	99.536
+	+		+	+	+	+		101291	315	396	17.54	99.30	99.61	99.69	99.533
+	+	+	+	+	+	+	+	101204	231	483	15.33	99.30	99.53	99.77	99.533
+	+		+	+	+	+	+	101286	320	401	15.88	99.29	99.61	99.69	99.530
	+					+		101235	303	452	25.85	99.26	99.56	99.70	99.506
+	+	+						101280	388	407	17.17	99.22	99.60	99.62	99.480
+								101125	778	562	19.48	98.69	99.45	99.24	99.126
					+			100054	350	1633	18.98	98.06	98.39	99.65	98.695
			+					100899	1323	788	12.44	97.95	99.23	98.71	98.627
	+							101044	1499	643	24.25	97.92	99.37	98.54	98.606
				+				99070	106	2617	3.92	97.32	97.43	99.89	98.199
		+						100251	1654	1436	1.37	97.01	98.59	98.38	97.988
						+		101312	6583	375	2.55	93.57	99.63	93.90	95.622
							+	94756	1178	6931	1.15	92.12	93.18	98.77	94.602

Заключение

В рамках исследования показана эффективность описанного подхода объединения результатов нескольких детекторов на базе алгоритма адаптивной кластеризации для повышения качества определения R-пигов. Из не решённых до конца проблем можно отметить довольно высокое значение MAE найденных положений R-пигов относительно истинных, которое хоть и понизилось для таких детекторов как Pan and Tompkins и Two

Moving Average, но всё же составляет порядка в среднем чуть меньше 20 отсчётов при частоте $f=360$ Гц. Для предотвращения этого недостатка целесообразно осуществлять локальное выравнивание положения R-пику, исследуя в рамках окрестности радиуса R вокруг найденного пика максимальное по модулю значение сигнала. В рамках описанной процедуры выравнивание можно осуществлять тремя разными способами: до кластеризации, после кластеризации и в оба момента времени.

Из дополнительных направлений исследования можно также отметить перспективность проведения экспериментов с большим числом отведений. Например, в МПТ-ВН также есть второе отведение, для которого аналогичным образом можно выявить наиболее качественные объединения детекторов, а затем – исследовать возможность повышения качества при совместном анализе обоих отведений сигнала ЭКГ одновременно.

Наконец, стоит отметить, что предложенный подход можно обобщить и на другие сигналы, для которых требуется решать задачу обнаружения аномалий в реальном времени по нескольким источникам, каждый из которых может ошибаться.

Библиографический список

1. Кузнецов А.А. Биофизика сердца. Методы обработки и анализа электрокардиографической информации при донозологических исследованиях: учеб. пособие / А.А. Кузнецов ; Владим. гос. ун-т имени Александра Григорьевича и Николая Григорьевича Столетовых. – Владимир: Изд-во ВлГУ, 2012. – 237 с.
2. Халайджи А.К. Анализ качества работы детекторов R-пику в сигнале ЭКГ в режиме реального времени / А.К. Халайджи // Сборник трудов ИСТ-2021, 2021.
3. Porr B. Popular ECG R peak detectors written in python (Version 1.0.2) / B. Porr // Zenodo, 2019. DOI: 10.5281/zenodo.3588108
4. Pan J. A Real-Time QRS Detection Algorithm / J. Pan, W.J. Tompkins // IEEE Transactions on Biomedical Engineering BME-32.3, 1985. – P. 230–236. DOI: 10.1109/TBME.1985.325532
5. Hamilton P.S. Open Source ECG Analysis Software Documentation / P.S. Hamilton // Computers in Cardiology, 2002. – V. 29. – P. 101-104. DOI: 10.1109/CIC.2002.1166717
6. Christov I.I. Real time electrocardiogram QRS detection using combined adaptive threshold / I.I. Christov // BioMedical Engineering OnLine, 2004. – V. 3. – №28. – P. 9. DOI: 10.1186/1475-925X-3-28
7. Lourenco A. Real Time Electrocardiogram Segmentation for Finger Based ECG Biometrics / A. Lourenco, H. Silva, P. Leite, R. Lourenco, A. Fred // BIOSIGNALS, 2012. – P. 49-54. DOI: 10.5220/0003777300490054
8. Kalidas V. Real-time QRS detector using Stationary Wavelet Transform for Automated ECG Analysis / V. Kalidas, L. Tamil // 2017 IEEE 17th International Conference on Bioinformatics and Bioengineering (BIBE), 2017. – P. 457-461. DOI: 10.1109/BIBE.2017.00-12
9. Elgendi M. Frequency Bands Effects on QRS Detection / M. Elgendi, M. Jonkman, F. De Boer // The 3rd International Conference on Bio-inspired Systems and Signal Processing (BIOSIGNALS2010), 2010. – V. 1. – P. 428-431. DOI: 10.5220/0002742704280431
10. Luz E.J.S. ECG-based heartbeat classification for arrhythmia detection: A survey / E.J.S. Luz, W.R. Schwartz, G. Camara-Chavez, D. Menotti // Computer Methods and Programs in Biomedicine, 2016. – V. 127 – P. 144-164. DOI: 10.1016/j.cmpb.2015.12.008
11. de Chazal P. Automatic classification of heartbeats using ECG morphology and heartbeat interval features / P. de Chazal, M. O'Dwyer, R.B. Reilly // IEEE Trans. Biomed. Eng. – 2004. – V. 51. – № 7. – P. 1196–1206. DOI: 10.1109/tbme.2004.827359

12. Аль-Салех Г.И. Сравнительный анализ методов обработки электрокардиографического сигнала / Г.И. Аль-Салех, А.А. Рунова, А.В. Писарева, А.П. Николаев // Физика и радиоэлектроника в медицине и экологии. Труды XIII Международной научной конференции с научной молодёжной школой имени И.Н. Спиридонова. В 2-х книгах, 2018, Владимир-Суздаль. – Изд-во ООО «Графика». – С. 120-125.
13. Wang H. Optimal k-means Clustering in One Dimension by Dynamic Programming Ckmeans.1d.dp / H. Wang, M. Song. // The R Journal. – 2011. – V. 3. – № 2. – P. 29–33. DOI: 10.32614/RJ-2011-015

A.K. Khalaydzh

ALGORITHM FOR DETECTING R-PEAKS OF THE ECG SIGNAL BASED ON CLUSTERING THE RESULTS OF INDIVIDUAL DETECTORS

Bauman Moscow state technical university, Moscow

Purpose: The main goal of this paper is to increase the quality of R-peak detection procedure in real-time operation of the ECG signal, containing arrhythmias. Innovation of this research is in using adaptive 1D clustering technique to combine the results of several individual detectors that analyze ECG signal simultaneously.

Design / methodology / approach: The paper describes, how to combine the results of individual R-peak detectors. The author provides an explanation, why usual greedy approach doesn't provide optimal solution for this problem and suggests using merge sort to flatten all samples from all detectors on the uniform axis and clusterization procedure to combine these samples into clusters, interpreting their centers as final R-peak positions. To make it possible the author suggests to use optimal clusterization algorithm KMeans1D, having dynamic programming structure, and provides a heuristic rule for early stopping and automated adaptive determination of an optimal clusters count. The quality of the proposed approach is measured by ECG signals from MIT-BIH according to the inter-patient division paradigm.

Findings: As a result, paper shows that final detectors that obtain the best quality indeed use several detectors. Moreover, there is the best configuration for train, test subsets and both of them, that contain four detectors: Pan and Tompkins detector, Two Moving Average detector, Christov detector and DumbMax-detector with threshold 50.

Research limitations/implications: The research analyzes signals from the only database – MIT-BIH, because the results of the overall research were used further in the task of automated arrhythmia classification in real-time mode. Also the author concludes two more perspective tasks for the further analysis and improvement of the proposed algorithm. First is about adjustment of the final positions of R-peaks that is important for the extraction morphological features or determination RR-intervals durations. According to proposed algorithm, that approach can be done before, after clusterization or at both steps. The second task is about using the proposed algorithm to make similar research of combining information from several ECG signal leads simultaneously instead of using the only one (MLII).

Originality/value: The value of the whole paper is the new approach of real-time detection of R-peaks that proposes using existing individual detectors, but obtains more robust and accurate results. The proposed algorithm can be used in similar tasks of real-time anomalies detection using several not-accurate individual detectors.

Keywords: one-dimensional adaptive clusterization, KMeans, R-peak detector, MIT-BIH

АНАЛИЗ СОВРЕМЕННЫХ МЕТОДОВОБРАБОТКИ ИЗОБРАЖЕНИЙ ДЛЯ УСТРАНЕНИЯ ШУМОВ И ИСКАЖЕНИЙ НА МЕДИЦИНСКИХ СНИМКАХ

Нижегородский государственный технический университет им Р.Е. Алексеева

В работе рассматривается проблема наличия шумов и искажений на медицинских снимках. Проводится обзор методов обработки изображений и способов калибровки камер и микроскопов. Также в работе приведен анализ применимости методов обработки в разных сферах медицины.

Ключевые слова: обработка изображений, калибровка камер, искажения на изображениях, медицинские снимки.

Постановка задачи

В настоящее время телемедицина является одной из наиболее растущих и востребованных сфер здравоохранения. Современные информационно-коммуникационные технологии позволяют широко внедрять телемедицину в практику лечебных учреждений, осуществлять обмен необходимой медицинской информацией с целью проведения консультаций, уточнения диагноза, дистанционной корректировки лечения, наблюдения за пациентами, а также профилактики заболеваний. Особенно актуально использование возможностей телемедицины для удаленных территорий с отсутствием квалифицированных врачей и ограниченными медицинскими ресурсами, а также в условиях эпидемий и самоизоляции.

В отдельную, имеющую большое практическое значение и широкие перспективы внедрения, можно отнести область телемедицины, связанной с передачей, обработкой и анализом различных цифровых медицинских изображений – в области радиологии, патологии тканей, клеток, дерматологии и т.д. При обработке цифровых изображений необходимо решать задачи повышения качества изображения, устранения искажений геометрии и цвета, восстановления поврежденных изображений, распознавания объектов на изображениях. Часто исследуемый объект не укладывается на одном фото и необходимо делать несколько снимков, а затем их «склеивать» для дальнейшей обработки и анализа. В области телемедицины особенно актуально качество передаваемых данных, при этом недопустимы изображения с такими дефектами, как:

- искажение цветопередачи: насыщенности, контрастности, цветового тона;
- искажение размеров и геометрии снимаемых объектов;
- размытость объектов.

Для повышения точности диагностирования такого рода искажения необходимо убирать, т.е. необходимо автоматизировать процесс обработки изображений с целью повышения качества передаваемой по каналам связи информации. Таким образом, необходимо уже на этапе получения снимков при помощи алгоритмов обработки изображений и методов калибровки устранить всевозможные искажения.

Задача обработки изображений для устранения искажений является актуальной при построении систем компьютерного зрения – решения задач обработки, анализа и классификации объектов на изображении. В медицинской области работ, связанных с подготовкой качественных цветных медицинских изображений для дальнейшей их обработки практически нет, это молодая и очень перспективная область исследования, требующая разработки новых решений. Для успешного решения задач в этой области необходимо провести анализ существующих алгоритмов обработки изображений, выявить возможность их адаптации к устранению искажений с помощью калибровки камер.

Анализ методов улучшения цифровых изображений

Часто на изображениях присутствуют различные дефекты: излишняя зашумленность, размытые границы, различные артефакты.

Резкость изображения – это уровень размытости границы между соседними участками изображения с различной оптической плотностью. Резкость является одним из основных атрибутов изображения, который наиболее важен для его анализа.

Улучшение резкости изображения возможно с помощью метода нечеткой логики, нейро-нечетких систем, нерезкой маскировки, пороговых преобразований.

Алгоритм нечеткой логики состоит из трех основных этапов: фаззификация изображения, модификация значений принадлежности и, при необходимости, дефаззификация изображения. После фаззификации, то есть преобразования данных изображения из четких сигналов в нечеткие множества, соответствующие нечеткие методы изменяют значения принадлежности.

Нейро-нечеткие системы – это системы, в которых нечеткая логика и множества корректируются с помощью методов нейронной сети. Нейронная сеть помогает обнаружить различные типы шума: соль и перец, гауссовский и негауссовский шум. Нечеткая логика с помощью некоторого фильтра, выбранного в зависимости от типа шума, устраняет его. Таким образом нейронные сети в данной системе используются для идентификации шума по статистическим параметрам, а нечеткая логика – для его удаления [1].

В подходе нерезкой маскировки улучшение изображения происходит за счет объединения отфильтрованного изображения высоких частот и исходного. В этом методе нерезкая маскировка применяется частичным способом для обнаружения краев и граничных линий на изображении. Затем с помощью операции сглаживания происходит удаление нежелательных краев на выбранных областях, которые представляют собой шум соли и перца. Изображение без шума добавляется к сглаженному изображению. Таким образом получается исходное изображение с уменьшенным шумом. Данный метод имеет два недостатка. Во-первых, он может усилить шум, присутствующий в изображении. Во-вторых, он усиливает резкие переходы.

Пороговые преобразования особенно полезны для сегментации. Порог изображения – это процесс отделения объектов на изображении от его фона. Пороговые значения можно разделить на две основные категории: глобальные и локальные. Методы глобального порогового значения выбирают одно пороговое значение для всего изображения. Это считается операцией точечной обработки. Локальное адаптивное пороговое значение использует различные значения для каждого пикселя в соответствии с информацией о локальной области. Локальные методы пороговой обработки применяются к изображениям, имеющими неоднородную фоновую подсветку или сложные фоны. Это связано с тем, что гистограмма таких изображений дает более двух пиков, что затрудняет глобальную технику пороговой обработки для отделения объектов от фона, таким образом, локальные методы пороговой обработки являются предпочтительными.

Кроме размытых границ изображение может содержать различные шумы. Шум – это дефект изображения, при котором пиксели разного цвета хаотично расположены на фото. Эти пиксели не соответствуют действительному, оригинальному изображению. Существуют различные виды шумов: «соль-перец», гауссовский шум, дробовой шум. Наиболее популярными алгоритмами для удаления данного дефекта являются линейный, адаптивный, медианный и средний фильтры и фильтр Калмана. Эффективность работы алгоритмов удаления шума оценивается по различным метрикам. Наиболее популярными являются Mean Square Error (MSE) и Root mean square error (RMSE).

Линейные фильтры используются для удаления определенных видов шумов [2]. Они имеют тенденцию размывать острые края, разрушать линии и другие мелкие детали изображения. В общем случае сглаживающий фильтр устанавливает для каждого пикселя среднее значение или средневзвешенное значение самого пикселя и его ближайших

соседей. Из-за этого линейные фильтры на практике используются для снижения шума. Также линейные фильтры часто используются в качестве основы для нелинейных фильтров.

Адаптивный фильтр – это саморегулирующаяся система, использующая для обработки рекурсивный алгоритм. Сначала сравниваются входные и обучающие сигналы и, соответственно, генерируется сигнал ошибки, который используется для корректировки некоторых ранее принятых параметров фильтра под действием входящего сигнала. Настройка параметров фильтра продолжается до установившегося состояния.

Средний фильтр – это простой пространственный фильтр. В нем используется скользящее окно. Алгоритм заменяет центральное значение в окне средним значением всех соседних значений пикселей, включая его самого. Этот процесс повторяется для всех пикселей на изображении.

Медианный фильтр – это простой и мощный нелинейный фильтр, основанный на статике порядка, отклик которого основан на ранжировании значений пикселей, содержащихся в области фильтра. Медианный фильтр также следует принципу движущегося окна, аналогично среднему фильтру. В этом фильтре нет замены значения пикселя изображения средним значением всех соседних пикселей. В нем производится замена медианным значением. Медианная фильтрация выполняется путем сортировки всех значений пикселей из окружающего окружения в числовом порядке, а затем замены рассматриваемого пикселя средним значением пикселя.

Фильтр Калмана работает через механизм предсказания и коррекции, из-за этого он также называется математической процедурой. Фильтр Калмана объединяет все доступные данные, то есть измеренные, системные знания и измерительные приборы, чтобы получить оценку искомых переменных таким образом, чтобы погрешность между измеренными и исходными данными была статистически минимизирована. Обычно этот фильтр используется для уменьшения белого шума.

Слабый контраст – наиболее распространенный дефект фотографических, сканерных и телевизионных изображений, обусловленный ограниченностью диапазона воспроизводимых яркостей. Под контрастом обычно понимают разность максимального и минимального значений яркости и представляют его в виде гистограммы яркости. Путем цифровой обработки контраст можно повысить, изменяя яркость каждого элемента изображения и увеличивая диапазон яркостей. Для этого разработано несколько методов.

Эквализация гистограммы – общий метод улучшения изображений. Он применяется не только для изменения контрастности, но и для удаления шума, редактирования резкости. Этот метод показывает свою эффективность при работе с одним каналом. Например, для черного-белого изображения можно построить гистограмму, определенную на множестве значений $[0; 2b_{ppr}]$, где b_{ppr} – количество бит. При проверке на темных изображениях большая часть нулевых значений будут сконцентрированы в конце множества значений. Для восприятия человека необходима примерно равная концентрация значений яркости на всем множестве. Эквализация гистограммы осуществляется линейно.

Адаптивное выравнивание гистограммы расширяет активный диапазон значений при сжатии гистограммы. Данный метод используется для улучшения различий в изображениях. Выравнивание гистограммы выделяется только на месте локального сравнения. Если изображение содержит области, которые значительно светлее и темнее, контраст в этих областях не будет в достаточной степени увеличен. Таким образом, адаптивное выравнивание гистограммы правильно вычисляет область изображения. Данный метод увеличивает контраст изображений за счет преобразования значений яркости изображения. Значение преобразования контраста вычисляется для каждой из этих областей независимо. Оптимальный размер области зависит от типа входного изображения и в превосходной степени определяется в ходе экспериментов. Преимущество этого метода заключается в том, что он быстр и прост. Результаты этого

метода показывают превосходство над обычно используемым выравниванием гистограммы.

Гомоморфная фильтрация – это обобщенная техника для цифровой обработки сигналов и изображений. Гомоморфная обработка выполняется для каждого поддиапазона отдельно. Известно, что гомоморфная обработка изображений выполняется в логарифмической области, которая преобразует изображение в компоненты освещения и отражения. Повышение коэффициента отражения усиливает детали изображения. Таким образом, применение этого процесса в каждом поддиапазоне улучшает детализацию изображения. Наконец, выполняется обратное аддитивное вейвлет-преобразование для гомоморфных расширенных поддиапазонов, чтобы получить инфракрасное изображение с лучшими визуальными деталями. Этот алгоритм применяется для улучшения контрастности инфракрасных изображений.

Методы математической морфологии [3]. Эти методы действуют на основе структурных свойств объектов. Они используют математические принципы и отношения между категориями для извлечения компонентов изображения, которые полезны при описании формы зон. Морфологические операторы нелинейны, и два набора данных являются их входными данными. Первый набор содержит исходное изображение, а второй описывает структурный элемент (маску). Исходное изображение является двоичным или представляет собой уровни серого, а маска представляет собой матрицу, содержащую ноль и единицу значений. Методы математической морфологии направлены, прежде всего, для извлечения некоторых свойств изображения, полезных для его представления и описания. Например, контуров, остовов, выпуклых оболочек.

Калибровка камер.

Некоторые искажения на изображениях могут быть вызваны не идеальностью оптической системы и физическими особенностями некоторых объективов и методов съемки. Один из видов оптического искажения это дисторсия, т.е. ошибка оптических систем, в результате которой нарушается геометрическое подобие между объектом и его изображением [4]. Это связано с тем, что коэффициент линейного увеличения изменяется по мере удаления отображаемых предметов от оптической оси. Существуют два основных вида дисторсии: радиальная дисторсия и тангенциальная дисторсия.

Радиальная дисторсия — искажение изображения в результате неидеальности параболической формы линзы. Искажения, вызванные радиальной дисторсией, равны 0 в оптическом центре сенсора и возрастают к краям. Как правило, радиальная дисторсия вносит наибольший вклад в искажение изображения.

Тангенциальная дисторсия — искажения изображения, вызванные погрешностями в установке линзы параллельно плоскости изображения.

Дисторсию можно определить в процессе калибровки камеры. Калибровка камеры – это задача, заключающаяся в поиске точных математических отношений между точками в 3D пространстве и их 2D проекциями на экране камеры.

Чтобы найти проекцию 3D-точки на плоскость изображения, сначала надо перевести точку из мировой системы координат в систему координат камеры. Затем точка проецируется на плоскость изображения и связывается с проекцией в пиксельных координатах.

Цель процесса калибровки — найти матрицы внешних и внутренних параметров, а также коэффициенты радиальных и тангенциальных искажений объектива. Калибровка камеры позволяет минимизировать дисторсию объектива.

Существуют различные способы решения задачи калибровки: калибровка с помощью тест-объекта и самокалибровка.

В калибровочных моделях машинного и компьютерного зрения используются эталонные сетки, калибровочная матрица определяется с использованием изображений известного массива точек объекта, например, шаблона шахматной доски. Это связано с

тем, что у шахматной доски есть определенные свойства, которые делают ее удобной для калибровки:

- Это плоскость. Отсутствует необходимость работы с «глубиной».
- На этой плоскости есть точки, которые можно однозначно идентифицировать.
- Можно легко извлечь эти точки.
- Эти точки физически лежат на прямых линиях, независимо от того, как они выглядят на камере.

Существуют различные алгоритмы калибровки с помощью тест-объекта: модель Цая, Heikkila&Silven, Zhang.

Модель калибровки Цая предполагает, что некоторые параметры камеры предоставляются изготовителем, чтобы уменьшить первоначальную догадку оценки. Он требует n признаков точек на изображении и решает проблему калибровки с набором из n линейных уравнений, основанных на ограничении радиального выравнивания. Используется модель радиального искажения второго порядка, в то время как не рассматриваются условия децентрализации искажения. Двухэтапный метод может справиться либо с одним изображением, либо с несколькими изображениями трехмерной или планарной калибровочной сетки, но координаты точки сетки должны быть известны.

Методика, разработанная Heikkila&Silven, сначала извлекает исходные оценки параметров камеры, а затем применяется нелинейная оценка наименьших квадратов для вычисления параметров искажения. Модель использует два коэффициента как для радиального, так и для децентрализованного искажения, а метод работает с одним или несколькими изображениями с 2D или 3D калибровочными сетками.

Метод калибровки Чжана требует, чтобы передняя панель была размещена в виде планарной сетки шахматной доски с разной ориентацией. Разработанный алгоритм использует извлеченные угловые точки шаблона шахматной доски для вычисления проективного преобразования между точками изображения n разных изображений с точностью до масштабного коэффициента. После этого внутренние и внешние параметры камеры восстанавливаются с использованием решения с закрытой формой, а термины радиального искажения третьего и пятого порядка восстанавливаются в линейном решении наименьших квадратов.

При самокалибровке не используется объект калибровки, а метрические свойства камеры и изображенной сцены восстанавливаются из набора «некалиброванных» изображений с использованием ограничений на параметры камеры или на изображенной сцене. Самокалибровка обычно применяется в 3D-моделировании для обновления проективной реконструкции до единицы, которая является метрической. В общем случае применяются три типа ограничений для выполнения самокалибровки: ограничения сцены, ограничения движения камеры или ограничения на встроенные параметры камеры.

Еще одно искажение с которым приходится сталкиваться это искажение цветопередачи. Цветопередача зависит от освещения, типа светочувствительного сенсора камеры, объектива, а также заводской калибровки камеры. Для получения более точной информации о цвете необходимо провести калибровку всей измерительной системы в текущих условиях.

Под цветовой калибровкой понимается преобразование цветовой схемы изображения, которое было получено с помощью калибруемой камеры, к реальным, видимым человеческим глазом, цветам, то есть к некоторым эталонным значениям.

Для калибровки цвета камеры необходимо представить цвет в виде трех значений RGB. Чтобы получить однородное представление цветов при определенной цветовой температуре освещения необходимо использовать профили цвета – данные о цветовом охвате камеры.

Решить задачу цветовой калибровки камеры на основе яркостно-цветовых признаков изображения необходимо, так как результатом качественной калибровки камеры является повышенная диагностическая ценность изображений.

Применение методов улучшения цифровых изображений в телемедицине.

Используя информационные технологии в телемедицине, не получится создать универсальный метод обработки, который подойдет для любой сферы телемедицины. Связано это с тем, что в одних областях важна резкость изображения, но неважен цвет. В других областях контрастность изображения является решающим фактором для того, чтобы поставить правильный диагноз. К примеру, рентгеновские снимки близки к бинарным, поэтому там не так важен контраст, но очень важна резкость. Чтобы повысить резкость, можно увеличить дозу рентгеновского излучения, но это опасно для здоровья. Поэтому хорошим решением будет снизить дозу излучения и как-то компенсировать потерю резкости. Для решения данной задачи можно использовать методы нечеткой логики, нерезкой маскировки и пороговых преобразований. В компьютерной томографии также важна резкость изображения. Кроме классических методов, для повышения резкости изображений используются нейросетевые алгоритмы. По началу алгоритмы данного класса использовались лишь как вспомогательные инструменты. Таким образом, метод нейро-нечетких систем подготавливает изображение для обработки методом нечеткой логики. С развитием нейросетевых технологий стали появляться новые способы обработки изображений. Одним из них является метод, основанный на использовании генеративной нейронной сети для увеличения разрешения изображений – SuperresolutionGAN [5]. Данная нейронная сеть позволяет на основе изображения с низким разрешением получать изображения с большим разрешением. Преимуществом данного способа является простота в использовании, по сравнению с калибровкой камеры или микроскопа. К недостаткам стоит отнести появление новых шумов на изображении. Чтобы понять перспективы Superresolution нейронной сети, необходимо разобрать на практике, на сколько критичны недостатки. На сегодняшний день существует несколько реализаций данной модели сети, в том числе и предобученные сети на основе медицинских снимков. Одним из инструментов, который поддерживает эту технологию, является продукт от компании Intel – OpenVINO.

Калибровка цвета является достаточно сложной задачей, так как трудно объективно оценить цветовую палитру, представленную на изображении. Но это не отменяет того факта, что калибровка цвета необходима в таких областях как: офтальмология, микроскопия. На основе правильно откалиброванной камеры по цвету и контрасту врач-офтальмолог может с большей уверенностью определить наличие поврежденных тонких сосудов, что позволит избежать дегенерацию желтого пятна. Решая задачу калибровки цвета никак не обойтись без человеческого фактора, так как конечный результат будет оценивать человек. Для того, чтобы как-то стандартизировать получаемые результаты, большинство методов по коррекции цвета основаны на использовании стандартизированных цветов.

Заключение

Задача устранения шумов и искажения решается уже не первый год. За все это время был накоплен большой базис по этой тематике. Но задача все еще остается актуальной, так как множество практических вопросов не решено. Одни из методов работает недостаточно эффективно, другие сложны в реализации. С развитием информационных технологий у нас появляется больше возможностей для решений поставленной задачи. Но пока сложно сказать, какой из этих методов позволит получить необходимое качество изображения. Для этого нужно провести научные исследования и оценить качество работы методы. Также стоит отметить, что задача устранения шумов и искажений изображений, так же задача калибровки камер и микроскопов относится не

только в сфере телемедицины. Но данная область накладывает свои требования к создаваемым системам для обработки изображений. Поэтому общая задача обработки медицинских снимков разбивается на подзадачи, в зависимости от направления, будь то офтальмология или рентгеновские снимки. В качестве способов обработки и калибровки можно использовать классические алгоритмы, но не стоит забывать про методы, основанные на нейросетевых технологиях, так как они достаточно перспективны. Но перед тем, как их использовать, необходимо провести исследования и убедиться, что вносимые искажения легко поправимы или не критичны для конечного результата.

Библиографический список

1. Shihabudheen K.V., Pillai G.N. “Recent advances in neuro-fuzzy system: A survey”, Knowledge-Based Systems. 2018.
2. Ishchenko A., Nesteryuk A., Polyakova M. “The Technique of Extraction Text Areas on Scanned Document Image Using Linear Filtration”, Applied Aspects of Information Technology. 2019.
3. Mishra M., Panigrahi R.R., Rout P.K. “A combined mathematical morphology and extreme learning machine techniques-based approach to micro-grid protection”, Ain Shams Engineering Journal. 2019.
4. J Francis., Bian L. “Deep Learning for Distortion Prediction in Laser-Based Additive Manufacturing using Big Data”, Manufacturing Letters. 2019.
5. Super-resolution imaging [Электронный ресурс]: Материал из Википедии – свободной энциклопедии. Режим доступа: https://en.wikipedia.org/wiki/Super-resolution_imaging

D.A. Koblyakov, A.Y. Myagkov, E.A. Gusev, M.S. Yunochkin

OVERVIEW OF IMAGE PROCESSING TECHNIQUES FOR ELIMINATING NOISE AND DISTORTION IN MEDICAL IMAGES

Nizhny Novgorod State Technical University n.a. R.E. Alekseev

The paper deals with the problem of the presence of noise and distortion in medical images. An overview of image processing methods and calibration methods for cameras and microscopes is provided. The paper also provides an analysis of the applicability of treatment methods in various fields of medicine.

Keywords: image processing, camera calibration, image distortion, medical images.

АЛГОРИТМЫ КОМПЬЮТЕРНОГО ЗРЕНИЯ В ЗАДАЧЕ РАСПОЗНАВАНИЯ ЧАСТИЦ ВЕЩЕСТВА НА МИКРОСКОПИЧЕСКИХ СНИМКАХ

Нижегородский государственный технический университет им. Р.Е. Алексеева

В работе исследуется применение алгоритмов компьютерного зрения в задаче обработки микроскопических снимков. Разработано программное обеспечение для распознавания частиц вещества на изображениях и определения характеристик найденных объектов. Рассчитаны показатели оценки качества распознавания и представлены примеры использования разработанной программы.

Ключевые слова: компьютерное зрение, разработка программного обеспечения, микроскопия.

Введение

Определение размеров и формы частиц вещества при изготовлении порошков, эмульсий, суспензий является важной задачей, так как указанные параметры позволяют оценить свойства и качество изготовленного образца. Одним из наиболее широко применяемых методов определения размеров частиц вещества является микроскопия. Автоматизация анализа изображений, полученных при микроскопии, позволит сократить время работы, увеличить производительность труда, а также уменьшить влияние человеческого фактора и вероятность ошибок.

Исследование применения алгоритмов компьютерного зрения для распознавания частиц вещества на снимках, полученных при микроскопии

В ходе данного исследования была разработана система для распознавания частиц вещества на микроскопических снимках с использованием алгоритмов компьютерного зрения. Примеры анализируемых изображений представлены на рис. 1.

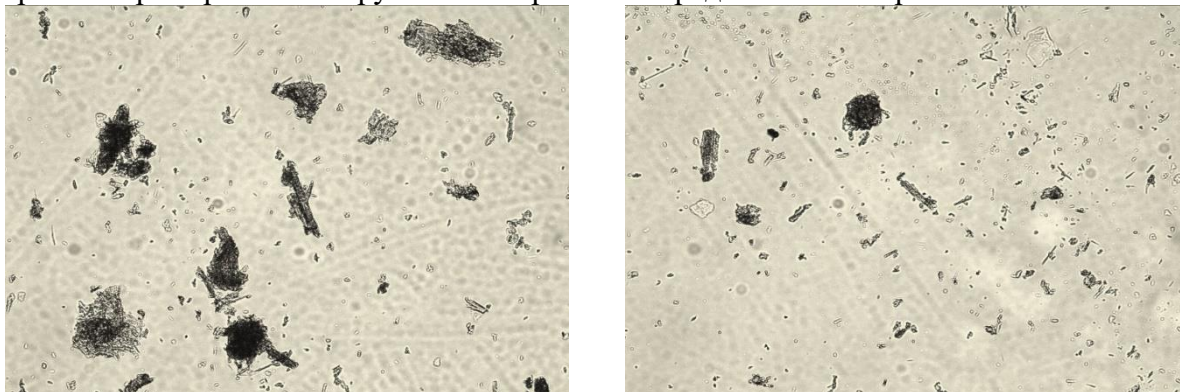


Рис.1. Примеры анализируемых изображений

Разработанная система позволяет обнаружить частицы вещества на изображениях, а также рассчитать набор параметров, характеризующих их размеры.

Способ определения размера частицы зависит от ее формы. Для сферических частиц размер определяется диаметром. Для частиц неправильной формы используются другие параметры, такие как диаметр Фере, а также ширина и длина частицы (рис. 2). Диаметр Фере рассчитывается как расстояние между двумя касательными, проведенными к противоположным сторонам частицы параллельно заданному направлению, одинакового для всего ряда частиц.



Рис. 2. Характеристики размеров частиц неправильной формы

Поскольку частицы правильной геометрической формы встречаются достаточно редко, при разработке программного обеспечения реализовано определение таких параметров как диаметр Фере, длина и ширина каждой частицы, а также выполнено построение графиков распределения частиц по значениям указанных параметров.

Для разработки системы был выбран язык C++, в качестве библиотеки компьютерного зрения выбрана библиотека OpenVino [1].

В реализованном в работе решении задачи распознавания частиц можно выделить следующие основные этапы: фильтрация изображения, бинаризация и применение морфологических операций, поиск контуров объектов (частиц), расчет параметров, характеризующих размеры частиц. Поскольку от правильного выбора алгоритма бинаризации во многом зависит качество распознавания, в ходе проведения исследования были апробированы несколько алгоритмов бинаризации, такие как, пороговая бинаризация, алгоритм Оцу [2], алгоритм Зака [3] и алгоритм адаптивной бинаризации (где пороговое значение для каждого пикселя рассчитывается как среднее значение по окрестности данного пикселя). Наилучшие результаты показали алгоритмы бинаризации Оцу и Зака. Отсутствие необходимости подбора оптимального значения порога бинаризации является достоинствами указанных алгоритмов. Экспериментально установлено, применение алгоритма Оцу позволяет отсеять объекты, не предназначенные для анализа, но при этом существенно уменьшает размеры искомых объектов. Таким образом, для анализа микроскопических снимков был выбран алгоритм бинаризации Зака.

Для устранения мелких объектов, не являющихся искомыми, в работе использована фильтрация изображения. После апробации алгоритмов фильтрации были выбраны: фильтр Гаусса, морфологические операции сужение и расширение, и медианный фильтр. Отметим, что фильтр Гаусса применяется для исходного изображения, а морфологические операции и медианный фильтр – для бинаризованного.

Приведем описание реализации всех описанных выше этапов обработки изображения и укажем настроенные значения параметров для каждого использованного в работе алгоритма (значения параметров были подобраны эмпирически).

Изображение загружается в формате png, jpg, bmp или tif (рисунок 3а). Затем к снимку в формате cv::Mat применяется фильтр Гаусса (используется функция cv::GaussianBlur библиотеки OpenVino) с размером ядра 5x5, $\sigma_X = 15$, $\sigma_Y = 0$.

После загрузки изображения и применения фильтра Гаусса выполняется его преобразование в изображение в оттенках серого и осветление с помощью метода convertTo. Затем выполняется бинаризация алгоритмом Зака (функция cv::threshold с параметром THRESH_TRIANGLE). После этого для бинаризованного изображения (рисунок 3б) применяются морфологические операции – расширение (функция cv::dilate)

и сужение (функция `cv::erode`). Размер ядра операции расширения равен 10×10 пикселей, операции сужения – 5×5 пикселей. Далее к полученному изображению применяется медианный фильтр с размером фильтра 9×9 (рисунок 3в).

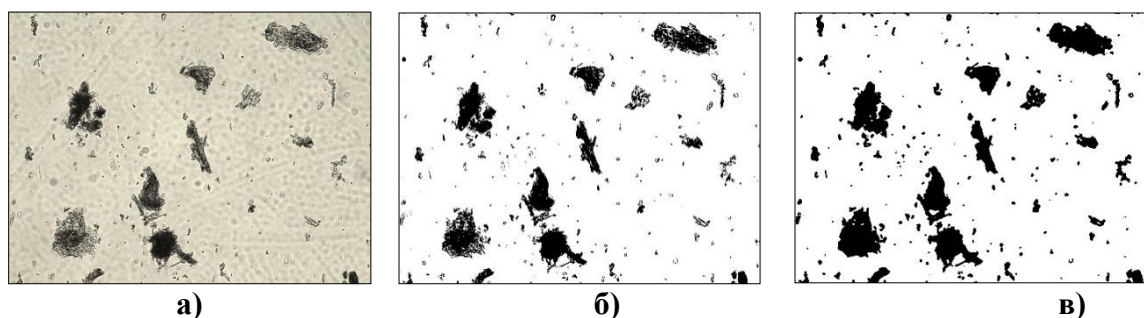


Рис. 3.

- а) исходный снимок;
- б) снимок после применения алгоритма бинаризации Зака;
- в) снимок с применением морфологических операций и медианного фильтра

С помощью функции `cv::findContours` на полученном изображении выполняется поиск контуров частиц.

Для того, чтобы рассчитать и построить график распределения частиц по значениям диаметра Фере, для каждой частицы выполняется поиск контура ограничивающего прямого прямоугольника (функция `cv::boundingRect`) (рисунок 4а).

Аналогично, чтобы рассчитать и построить график распределения частиц по значениям ширины и длины, для каждой частицы выполняется поиск контура ограничивающего вращающегося прямоугольника (функция `minAreaRect`) (рисунок 4б).

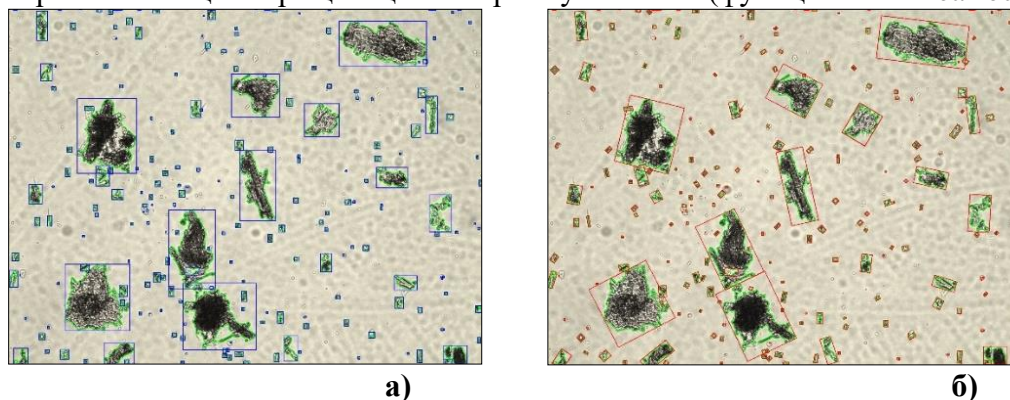


Рис. 4.

- а) снимок с выделенными ограничивающими прямыми прямоугольниками;
- б) снимок с выделенными ограничивающими вращающимися прямоугольниками

Оценка качества разработанной системы была проведена на базе данных изображений фармацевтических препаратов. В качестве показатели оценки эффективности разрабатываемой системы были использованы точность (`precision`), полнота (`recall`) и F-мера (`F-measure`). Значение показателя точности составило 76,1%, полноты – 94,3%, F-мера - 84,1%.

Для системы был разработан графический интерфейс пользователя (рисунок 5). Первоначально, пользователю необходимо загрузить снимок. До загрузки снимка невозможно увидеть статистику распределения размеров частиц или сохранить данные в файл.

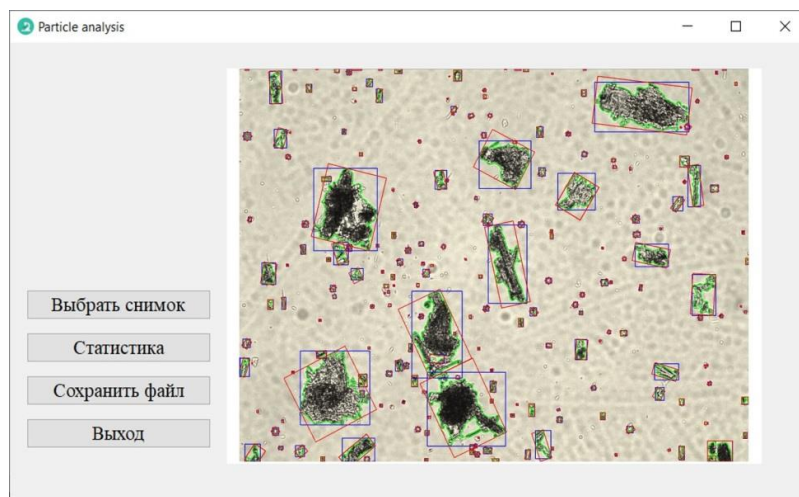


Рис. 5. Графический интерфейс программы

После выбора снимка выполняется его обработка, результатом которой является исходный снимок с размеченными на нем контурами частиц, а также контурами прямых прямоугольников и вращаемых прямоугольников (разметка происходит для каждой частицы). Пользователь имеет возможность построить графики распределения частиц по значениям горизонтального и вертикального диаметра Фере (рисунок 6). При этом по оси абсцисс отмечается значение диаметра Фере (вертикального или горизонтального). По оси ординат отмечается количество частиц с рассчитанным значением диаметра.

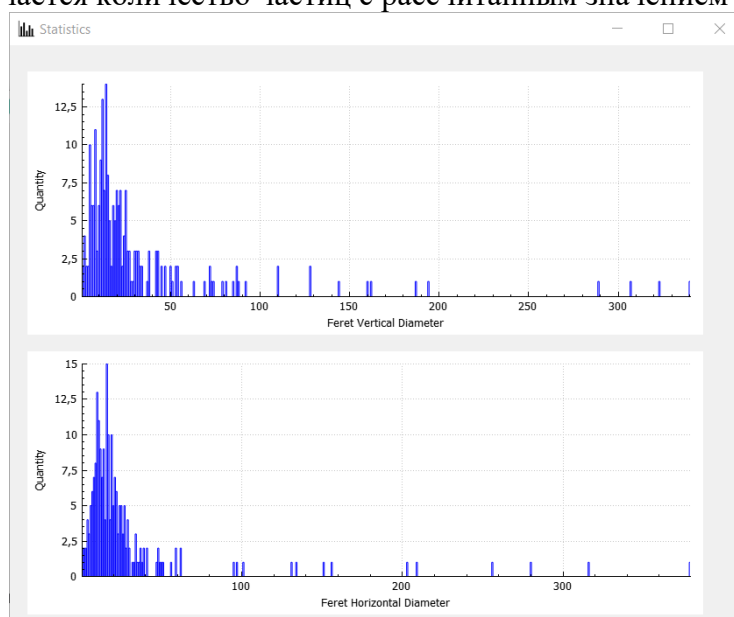


Рис. 6. Окно вывода графиков распределения частиц по значениям вертикального и горизонтального диаметров Фере

Все полученные данные можно сохранить в файл формата csv, который может быть открыт в программе Microsoft Excel в виде электронной таблицы.

Выводы

В ходе проведенного исследования изучены возможности применения алгоритмов компьютерного зрения в задаче распознавания частиц вещества на микроскопических снимках. Апробированы алгоритмы фильтрации и бинаризации, выбраны алгоритмы, показавшие наилучшие результаты для решения поставленной задачи. Выполнена реализация расчета таких характеристик размера частиц, как, длина частицы, ширина частицы и диаметр Фере. Рассчитаны значения показателей точности, полноты и F-меры для оценки качества работы программы. Для удобства использования программы разработан интерфейс пользователя.

Библиографический список

1. Open vino-toolkit [электронный ресурс]. URL: <https://software.intel.com/content/www/us/en/develop/tools/opencv.html> (дата обращения: 01.02.2021)
2. Bangare S.L Reviewing Otsu's Method For Image Thresholding / S.L.Bangare, A. Dubal, P.S.Bangare, S.T. Patil // International Journal of Applied Engineering Research. – 2015. – 10(9). – P. 21777-21783
3. Zack G.W. Automatic measurement of sister chromatid exchange frequency / G.W. Zack, W.E. Rogers, S.A. Latt // Journal of Histochemistry&Cytochemistry. –1977. – 25(7). – P.741-753

A.Y. Zakharova, P.A. Shagalova, D.A. Koblyakov

COMPUTER VISION ALGORITHMS FOR RECOGNIZING PARTICLES IN MICROSCOPIC IMAGES

Nizhny Novgorod state technical university n.a. R.E. Alekseev

In this article investigates the application of computer vision algorithms to the problem of processing microscopic images. Software has been developed for recognizing particles in images and determining the characteristics of found objects. The indicators for assessing the quality of recognition are calculated and examples of using the developed program are presented.

Keywords: computer vision, software development, microscopy.

М. Е. Маркин, М. Е. Федосова, А. В. Чернышов

ПРАКТИЧЕСКИЕ АСПЕКТЫ ВЫБОРА АРХИТЕКТУРЫ СВЕРТОЧНОЙ НЕЙРОННОЙ СЕТИ ДЛЯ БИНАРНОЙ КЛАССИФИКАЦИИ СНИМКОВ ФЛГ

Нижегородский государственный технический университет им. Р. Е. Алексеева

Для распознавания образов используют различные архитектуры нейронных сетей. Каждая архитектура наиболее удачно распознает один из видов образов. В работе рассматриваются практические аспекты использования различных архитектур сверточных нейронных сетей для решения задачи бинарной классификации флюорограмм грудной клетки с низкой информативностью снимков. Удовлетворительная точность достигнута при использовании нейронных сетей с архитектурой VGG и AlexNet.

Ключевые слова: искусственный интеллект, сверточная нейронная сеть, классификация изображений, классификация флюорограмм.

Введение

Исследования, направленные на решение задачи распознавания образов, начали активно развиваться в конце 20 века. Основной предпосылкой этих исследований стала разработка в 1957 году перцептрона Розенблата, который представлял собой одну из первых моделей запоминания и организации информации, реализуемых мозгом. Перцептрон Розенблата не был первой попыткой смоделировать деятельность человеческого мозга. Ранее над решением подобной задачи трудились Р. Буш и Ф. Мостеллер, Макколлок и Питтс, Фэрли и Кларк, Р. Эшби. На сегодняшний день задача распознавания образов понимается в гораздо более широком смысле, по сравнению с тем, как это трактовалось изначально. Под образом в настоящее время понимаются конкретные - зрительные и слуховые образы и абстрактные - опознание довода, обдумывание ситуации, нахождение решения задачи.

Для распознавания образов разработаны различные архитектуры нейронных сетей. Каждая архитектура наиболее удачно распознает один из видов образов. Так, например, для распознавания зрительных образов, широко применяются сверточные нейронные сети (Convolutional Neural Network, CNN). Такую архитектуру предложил в 1989 году Ян ЛеКун (LeNet) [1]. Идея сверточной сети LeNet заключается в чередовании сверточных и субдискретизирующих слоев, сеть принципиально многослойная. Архитектура сверточной сети LeNet отлично подходит для распознавания изображений. На основе данной сети разработаны различные модификации, среди которых можно выделить: AlexNet [2], ResNet [3], DenseNet [4], VGG [5]. Каждая модификация была разработана для конкретной задачи и имеет определенное количество слоев и число выделяемых признаков.

Сеть AlexNet схожа с сетью LeNet, однако имеет большее количество фильтров на каждом слое и большее количество слоев. Сеть состоит из 5 сверточных и 3 полносвязных слоев. В качестве функции активации сеть использует ReLU.

Первый сверточный слой фильтрует входное изображение размером 224x224x3 96 ядрами размером 11x11x3 с шагом 4 пикселя. Затем следует слой макспулинга и нормализации. Далее, на втором сверточном слое происходит фильтрация 256 ядрами размером 5x5x48. На третьем и четвертом сверточных слоях происходит фильтрация 384 ядрами размером 3x3x192. Пятый сверточный слой фильтрует 256 ядрами размером 3x3x192. После 5 сверточного слоя идут полносвязные слои размером 4096. Сеть насчитывает 62 миллиона обучаемых переменных.

Архитектура сети Alexnet приведена на рис. 1.

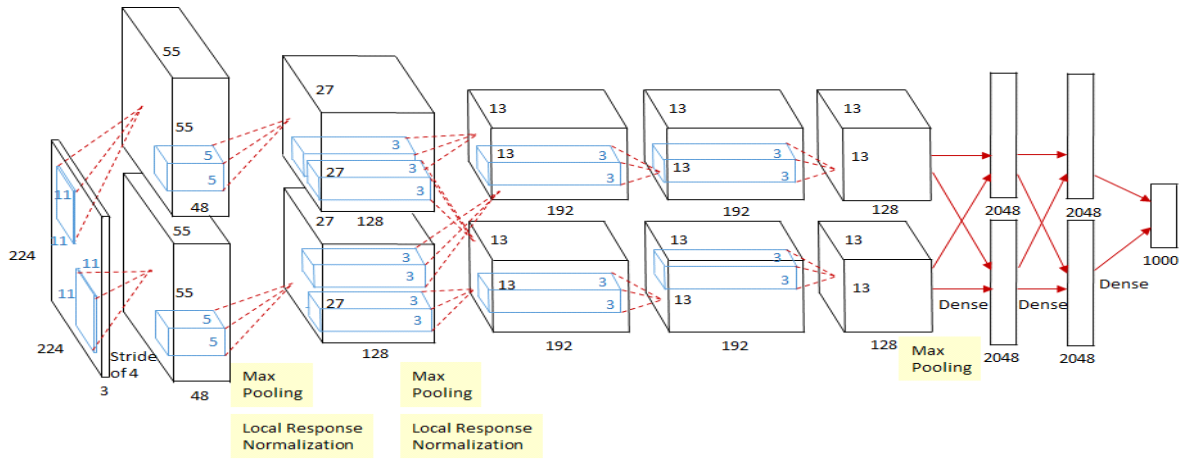


Рис. 1. Архитектура сети AlexNet

Сеть VGG является улучшенной версией сети AlexNet. Она была предложена К.Simonyan и А.Zisserman [5]. В этой сети использовано несколько фильтров 3x3 вместо 11x11 и 5x5. Архитектура сети VGG приведена на рис. 2.

Архитектура VGG16

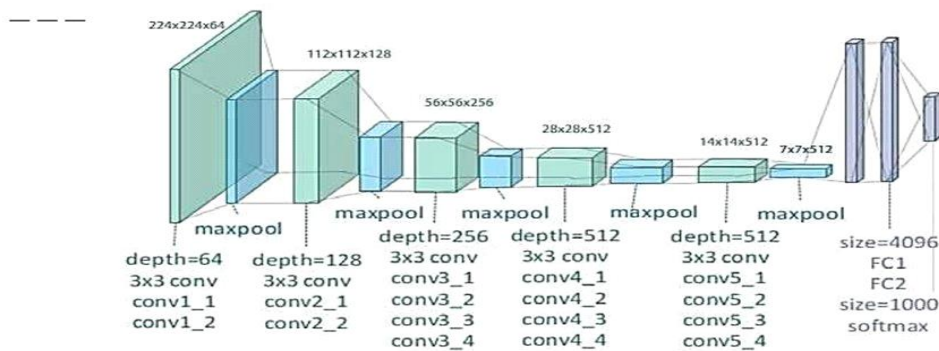


Рис. 2. Архитектура сети VGG

Сеть ResNet была разработана сотрудниками компании Microsoft: Kaiming He, Xiangyu Zhang, Shaoqing Ren, Jian Sun. Название сети – сокращенное от английского Residual Network, то есть остаточная нейронная сеть. ResNet – глубокая нейронная сеть с большим количеством слоев. Известная проблема сетей с большим количеством слоев – исчезающие градиенты [6,7]. До появления ResNet эта проблема решалась введением промежуточных слоев нормализации [8]. В ResNet для повышения точности использованы так называемые «быстрые соединения», которые пропускают один или несколько слоев. Разработаны сети ResNet с 18, 34, 50, 101, 152 слоями. При использовании 152 слоев получим 11 миллиардов обучаемых переменных. Архитектура сети ResNet приведена на рис. 3.

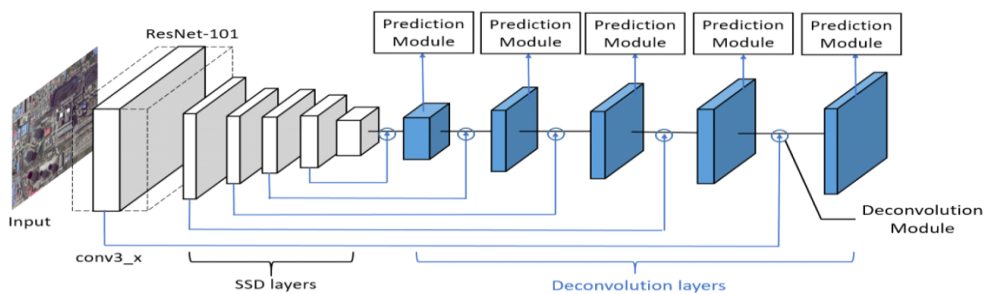


Рис. 3. Архитектура сети ResNet

Сеть DenseNet предложена командой разработчиков Gao Huang (Cornell University), Zhuang Liu (Tsinghua University), Laurens van der Maaten (Facebook AI Research) [4]. Эта сеть является модификацией сети ResNet. Сеть строится на том допущении, что многие слои ResNet могут быть случайно отброшены во время обучения, так как вносят очень небольшой вклад в передачу информации. Наличие таких слоев делает ResNet похожим на рекуррентные нейронные сети, но с числом параметров существенно большим, поскольку каждый слой имеет свои собственные веса. Архитектура DenseNet явно различает информацию, которая добавляется в сеть, и информацию, которая сохраняется. Слои DenseNet очень узки (например, 12 фильтров на слой). В DenseNet окончательное решение принимается на основе всех признаков сети за счет добавления только небольшого набора карт объектов и сохранению оставшихся карт признаков неизменными. Архитектура сети DenseNet приведена на рис. 4.

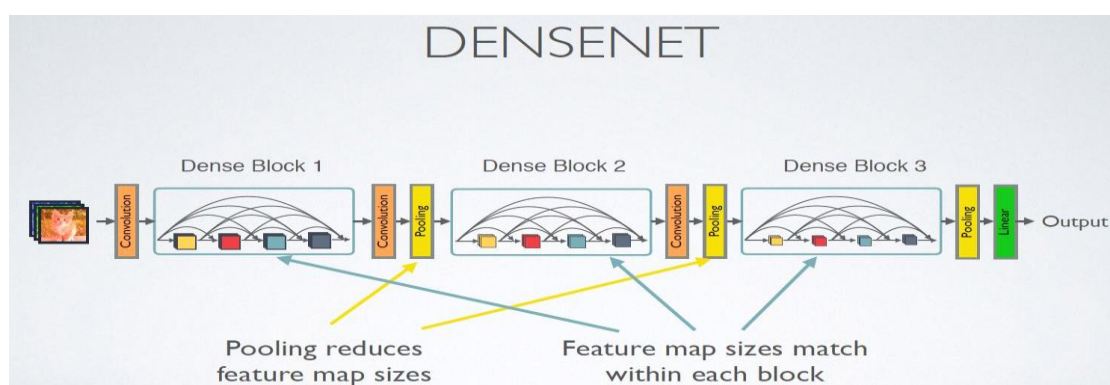


Рис. 4. Архитектура сети DenseNet

Использование нейронных сетей для классификации снимков ФЛГ

Использование методов искусственного интеллекта в медицине приобретает все большую распространенность. Одной из задач, которая с успехом может быть решена, является задача классификации легочных заболеваний. Традиционно для диагностирования большинства легочных заболеваний применяется флюорография (ФЛГ) и [рентген легких](#). По сути эти два метода схожи: ФЛГ является той же рентгеновской диагностикой, поскольку через исследуемые ткани проходят рентгеновские лучи. Однако принципиальная разница заключается в лучевой интенсивности и, соответственно, информативности. Информативность снимка ФЛГ меньше, так как меньше лучевая интенсивность.

ФЛГ является наиболее массовым методом профилактики наличия заболеваний легких. Обычно ФЛГ предшествует более глубокому рентгеновскому обследованию. Визуальную оценку снимка ФЛГ проводит врач по специальной методике: по полям от верхушки до нижнего края 2 ребра, от нижнего края 2 ребра до нижнего края 4, от нижнего 4 ребра и ниже. При этом оценивается состояние костной системы, симметричность полей правого и левого, легочный рисунок (сосуды, бронхи, интерстициальная ткань), проводится проверка на наличие очагов инфильтративных изменений, оцениваются корни, синусы, диафрагма, средостение и сердце, аорта, возрастные изменения. На этапе визуальной оценки врач определяет необходимость дообследования. По статистике, из 100 человек, направленных на дообследование, патология подтверждается не более чем у 10-15%. В период массовых обследований на врачей ложится колоссальная нагрузка, цена ошибки возрастает.

Как отмечалось выше, для автоматической классификации зрительных образов может использоваться аппарат сверточных нейронных сетей. Для успешного применения должна быть в первую очередь удачно выбрана архитектура сверточной сети и ее параметры. Объем тренировочных данных должен быть достаточным для обучения сети.

К тренировочным данным, кроме объема, предъявляются дополнительные требования: данные должны содержать все признаки, по которым будет происходить классификация.

Метрики качества работы нейронной сети

Качество работы нейронной сети оценивают на наборе данных, который не участвовал в обучении. Исходные данные для оценки берутся из матрицы ошибок, которая представляет собой табличную визуализацию предсказаний модели в сравнении с метками достоверности. Каждая строка матрицы ошибок (табл. 1) представляет экземпляры в предсказанном классе, а каждый столбец представляет экземпляры в реальном классе.

Таблица 1

Матрица ошибок

Предсказанный класс	Реальный класс	
	<p><i>TP (True Positive - истинно-положительное решение):</i> прогноз совпал с реальностью.</p>	<p><i>FP (False Positive - ложноположительное решение):</i> ошибка 1-го рода, нейронная сеть предсказала положительный результат, а на самом деле он отрицательный</p>
<p><i>FN (False Negative - ложноотрицательное решение):</i> ошибка 2-го рода – нейронная сеть предсказала отрицательный результат, но на самом деле он положительный.</p>	<p><i>TN (True Negative - истинно-отрицательное решение):</i> результат отрицательный, прогноз совпал с реальностью.</p>	

Точность классификации (ТК). Точность классификации является самой простой метрикой и определяется по формуле:

$$TK = \frac{TP + TN}{TP + TN + FN + FP} \cdot 100\%$$

Точность (Т). Во многих задачах метрика точность классификации не всегда является хорошим показателем эффективности модели. Один из таких случаев - когда распределение классов не сбалансировано (один класс встречается чаще, чем другие). Поэтому, для несбалансированных данных применяют специфичные показатели для каждого класса. Точность является одной из таких метрик, которая определяется так:

$$T = \frac{TP}{TP + FP} \cdot 100\%$$

Чувствительность. Чувствительность - еще одна важная метрика, которая определяется как доля выборок из класса, которые правильно прогнозируются моделью:

$$TPR = \frac{TP}{TP + FN} \cdot 100\%$$

Метрика F1. Эта метрика сочетает в себе точность и чувствительность. Эта метрика является гармоническим средним значением упомянутых выше метрик и определяемым как:

$$F1 = \frac{2 \cdot T \cdot TRP}{T + TRP}$$

В областях, связанных с медициной и биологией может использоваться специальная метрика, которая называется Ложная Тревога и вычисляется по формуле:

$$FPR = \frac{FP}{TN + FP} \cdot 100\%$$

Постановка задачи

Имелся обезличенный сбалансированный набор флюорограмм в количестве 25 000 штук, который проклассифицирован как «патология» и «норма». Размер снимков 256x256. Снимки черно-белые. Задача заключалась в выборе оптимальной архитектуры нейронной сети для автоматической бинарной классификации. Задача решалась в несколько этапов.

На первом этапе проводится подготовка данных для обучения нейронной сети. В результате были исключены заведомо непригодные для исследования снимки: с боковыми проекциями, сильно затемненные или засвеченные, снимки с нарушенной центрацией, а также снимки людей с сильным искривлением позвоночника. Кроме того, были убраны ФЛГ пациентов с кардиостимулятором.

На втором этапе проводилось обучение нейронной сети на тренировочном наборе данных, который содержал в общей сложности 25 000 проклассифицированных снимков.

Для анализа разделили сети на 2 группы:

- Группа 1: AlexNet и VGG (сети без пропусков слоев).
- Группа 2: ResNet и DenseNet (сети, где часть слоев отбрасывается).

На третьем этапе проводилась проверка качества работы сетей на сбалансированном (равное количество снимков с патологией и без) тестовом наборе данных с использованием метрики «точность». Число тестовых изображений – 620.

Результаты исследования представлены в таблице 2.

Таблица 2.

Результаты исследования

	Группа 1		Группа 2	
	AlexNet	VGG	ResNet	DenseNet
Точность	80,3%	80,8%	76,6%	77,4%

Наивысшую точность удалось достигнуть при batch = 15 и скорости обучения lr=0,0001 на сетях Группы 1. Причем, точность диагностики с использованием архитектуры VGG несколько лучше, чем для AlexNet. Архитектуры сетей Группы 2 оказались менее подходящими для решения поставленной задачи.

Выводы

На основе проведенных исследований можно сделать следующие выводы:

- Наличие «быстрых соединений» в сетях Группы 2 привело к тому, что часть важной информации потерялась, и точность по сравнению с сетями Группы 1, стала ниже.
- Фильтрация ядром размера 3x3 более предпочтительна, чем фильтрация ядрами размера 5x5 и 11x11.
- Сеть с архитектурой VGG может успешно применяться для решения задачи классификации снимков ФЛГ.

Библиографический список

1. LeCun Y., Boser B., Denker J. S., Henderson D., Howard R. E., Hubbard W. and Jackel L. D. Backpropagation Applied to Handwritten Zip Code Recognition, Neural Computation, 1(4):541-551, Winter 1989.
2. Alex Krizhevsky, Ilya Sutskever, Geoffrey E. Hinton. ImageNet Classification with Deep Convolutional Neural Networks. [Электронный ресурс]. URL: <https://papers.nips.cc/paper/2012/file/c399862d3b9d6b76c8436e924a68c45b-Paper.pdf>
3. Kaiming He, Xiangyu Zhang, Shaoqing Ren, Jian Sun. Deep Residual Learning for Image Recognition. [Электронный ресурс]. URL: <https://arxiv.org/pdf/1512.03385.pdf>
4. Gao Huang, Zhuang Liu, Laurens van der Maaten. Densely Connected Convolutional Networks. [Электронный ресурс]. URL: <https://arxiv.org/pdf/1608.06993.pdf>
5. Karen Simonyan, Andrew Zisserman. Very Deep Convolutional Networks for Large-Scale Image Recognition. [Электронный ресурс]. URL: <https://arxiv.org/pdf/1409.1556.pdf>
6. Bengio Y., Simard P., and Frasconi P. Learning long-term dependencies with gradient descent is difficult. IEEE Transactions on Neural Networks, 5(2):157–166, 1994.
7. Glorot X. and Bengio Y. Understanding the difficulty of training deep feedforward neural networks. In AISTATS, 2010.
8. Ioffe S. and Szegedy C. Batch normalization: Accelerating deep network training by reducing internal covariate shift. In ICML, 2015.

Maxim E. Markin, Marina E. Fedosova, Andrey V. Chernyshov

PRACTICAL USES OF DIFFERENT CONVOLUTIONAL NEURAL NETWORK ARCHITECTURES FOR BINARY X-RAY FLUOROGRAM CLASSIFICATION

Nizhny Novgorod state technical university n.a. R.E. Alekseev

Various neural network architectures are used for pattern recognition. Each architecture is most successful in recognizing one of the types of images. The paper discusses the practical aspects of using various architectures of convolutional neural networks to solve the problem of binary classification of X-ray fluorogram with low information content of images. Satisfactory accuracy was achieved using neural networks with VGG and AlexNet architecture

Keywords: artificial intelligence, convolutional neural network, image classification, X-ray fluorogram classification

ПРЕОБРАЗОВАНИЯ И УЛУЧШЕНИЕ ИЗОБРАЖЕНИЙ ДЛЯ НЕЙРОННЫХ СЕТЕЙ

Нижегородский государственный технический университет им. Р.Е. Алексева

В данной статье были рассмотрены алгоритмы по анализу, работы и преобразованию изображений для последующего использования в нейронных сетях. Был применен алгоритм выравнивания яркости и контрастности для обработки изображений и получения лучшего результата. Показан алгоритм компенсации изменения положения микроскопа при выполнении контрольных фотографий.

Ключевые слова: преобразование изображений, улучшение контрастности, фрагментарный поиск, предобработка для нейронных сетей.

Введение

Сбор и классификация дефектов в стали является важной частью обучения нейронной сети, от качества проведенной работы зависит точность классификации дефекта, а от количества распознаваемых повреждений зависит точность настройки оборудования и скорость принятия решения. Поэтому важно создать классификатор, отвечающий требованиям универсальности, достоверности, актуальности и избыточности. Разных результатов распознавания удастся достичь, используя разные типы классификаторов, например, используя линейный классификатор с SGD-обучением или метод опорных векторов и другие. Применяя разные методы классификации дефектов, следует пред обрабатывать изображения для получения точных результатов. Один из инструментов пред обработки это выравнивание изображения. Другие способы предобработки, которые могут повлиять на выполняемую задачу это изменение яркости, контрастности и цветового пространства. В данной работе рассмотрим подробнее некоторые преобразования над изображениями.

Основная часть

Был загружен датасет состоящий из 7 классов и 4 фотографий. Загруженные изображения для анализа представлены на рис.1.

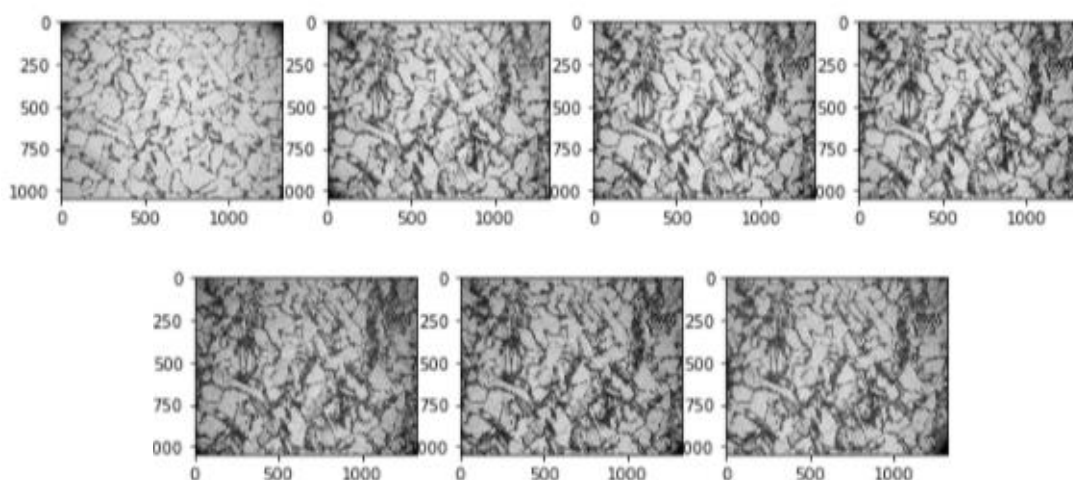


Рис. 1. Загруженные изображения

На рис.1 представлена серия фотографий, первая картинка на рисунке – это исходное состояние материала, последняя фотография – это наиболее поврежденное фото, промежуточные фотографии – это дефект в «развитии».

Важным для понимания качества изображения при работе и последующей обработке является график получаемый функцией `numpy.ravel`. Функция `ravel()` возвращает непрерывный и сжатый до одной оси массив. Полученные графики представлены на рис. 2.

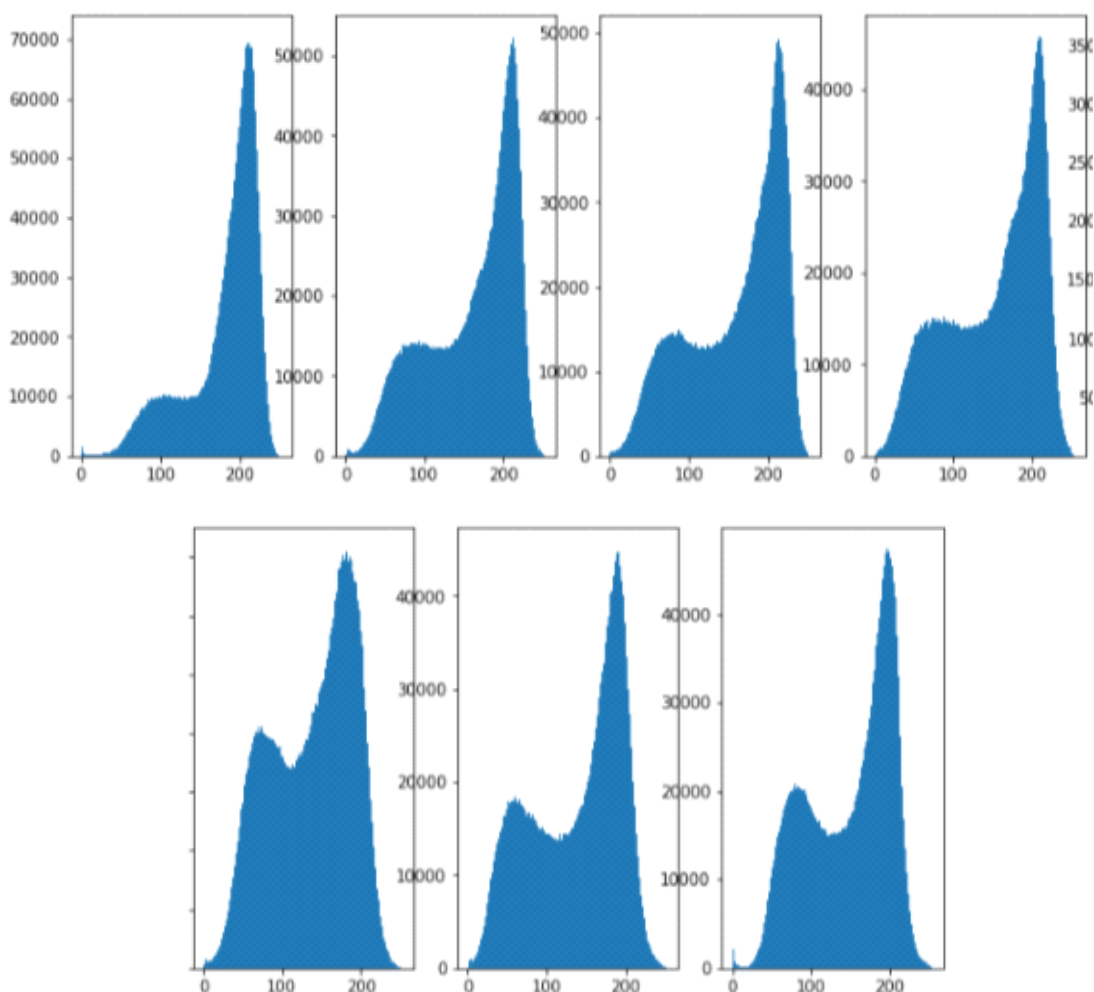


Рис. 2. Концентрация пикселей на изображении по яркости

Полученные графики показывают изменение количества пикселей и показывают накопление темных пикселей, а т.к. количество пикселей в изображении остается неизменным, можно сделать вывод, что количество белых пикселей уменьшается. Такое поведение можно объяснить накоплением маленьких трещин.

Далее были подсчитаны пиксели разных цветов. Основными были выбраны черные, белые и серые пиксели. По причине того, что исследуемые фотографии были сделаны на микроскопе, белые пиксели могут оказаться идеально белыми и немного не попадать в цвет. Аналогичная ситуация складывается с черными пикселями, поэтому для построения графиков и последующего анализа были взяты условно белые и условно черные точки. Белой точка будет считаться если значение пикселя в трех каналах (красный, синий и зеленый) будет больше 200, черный пиксель будет зарегистрирован если значение пикселя будет менее 100 и серым пиксель будет считаться при условии попадания в интервал от 120 до 130. Максимальное значение пикселя 255, минимальное 0.

Полученные результаты показывают изменение отношения состава пикселей на изображении. Серые пиксели показаны зеленым цветом, на графике видно, что количество серых пикселей остается практически неизменным. Белые пиксели показаны синим цветом, график показывает, что количество пикселей заметно изменяется и уменьшается. Черные пиксели на графике показаны оранжевым цветом, график

показывает, что количество пикселей увеличивается с ростом дефекта на изображении. Если посмотреть на изображение, то можно увидеть увеличение длины и ширины трещин, увеличения их количества. На фото каждый дефект показан черным цветом, поэтому изменение количества белых и черных пикселей говорит о увеличении дефектов на изображении, увеличении их площади и количества. Белые пиксели могут показать площадь не поврежденной поверхности.

Вычитание изображений

Далее рассматриваются результаты вычитания изображений. Для вычитания берется исходное изображение, из которого по пикселям вычитается последующее, затем все повторяется с следующей картинкой. Другой способ вычитания также на выходе получит разность изображений, но вычитаться всегда будет исходное изображение из остальных. На рис. 3, 4 показан результат вычитания и графики цветов.

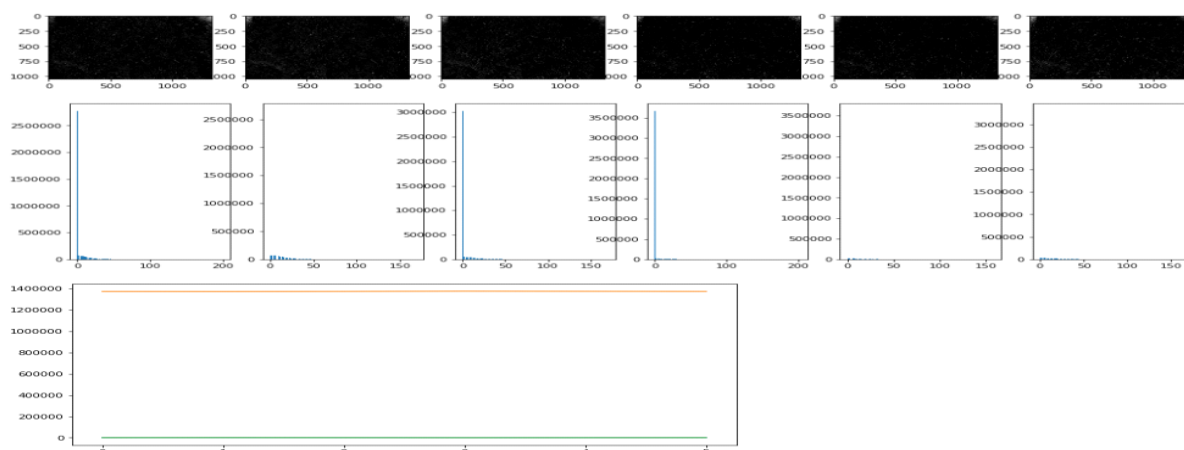


Рис. 3. Вычитание изображения, выбранного с последующим

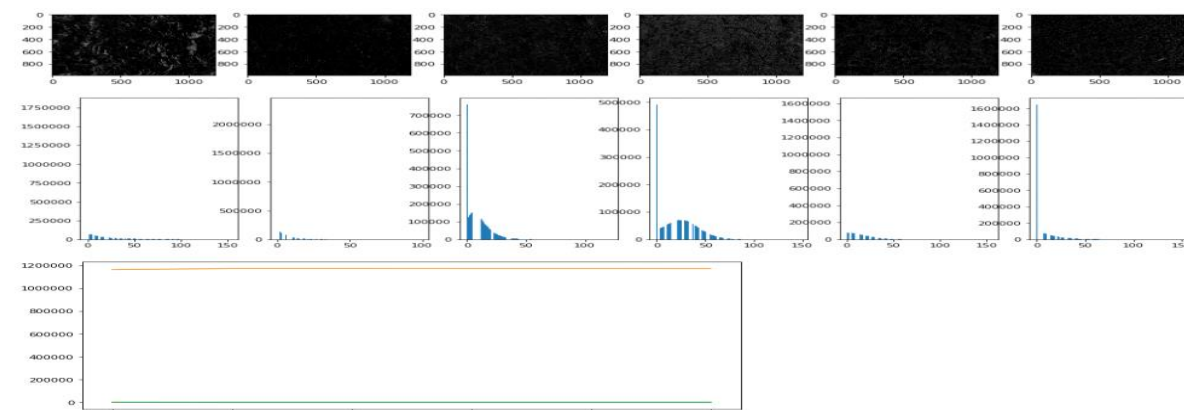


Рис. 4. Вычитание 1 изображения из остальных.

Полученный на рисунке 3 и 4 результат можно объяснить тем, что при съемке дефекта, микроскоп немного перемещается относительно изначальной фотографии и без коррекции перемещений сложно получить хороший результат.

Фрагментарный поиск.

Для минерализации изменения позиции микроскопа можно использовать фрагментарный поиск на изображении. На исходном изображении берется фрагмент картинки размером 300*300 пикселей. Далее эта картинка с использованием библиотеки OpenCV и метода cv2.matchTemplate определяется координата этого фрагмента. После получения координат этого фрагмента можно обрезать изображение и «подвинуть» изображение для более точного позиционирования и минимизировать изменение позиции микроскопа. На рис. 5 показан процесс выбора картинки и поиска ее на изображении. На рисунке также указаны координаты найденной области.

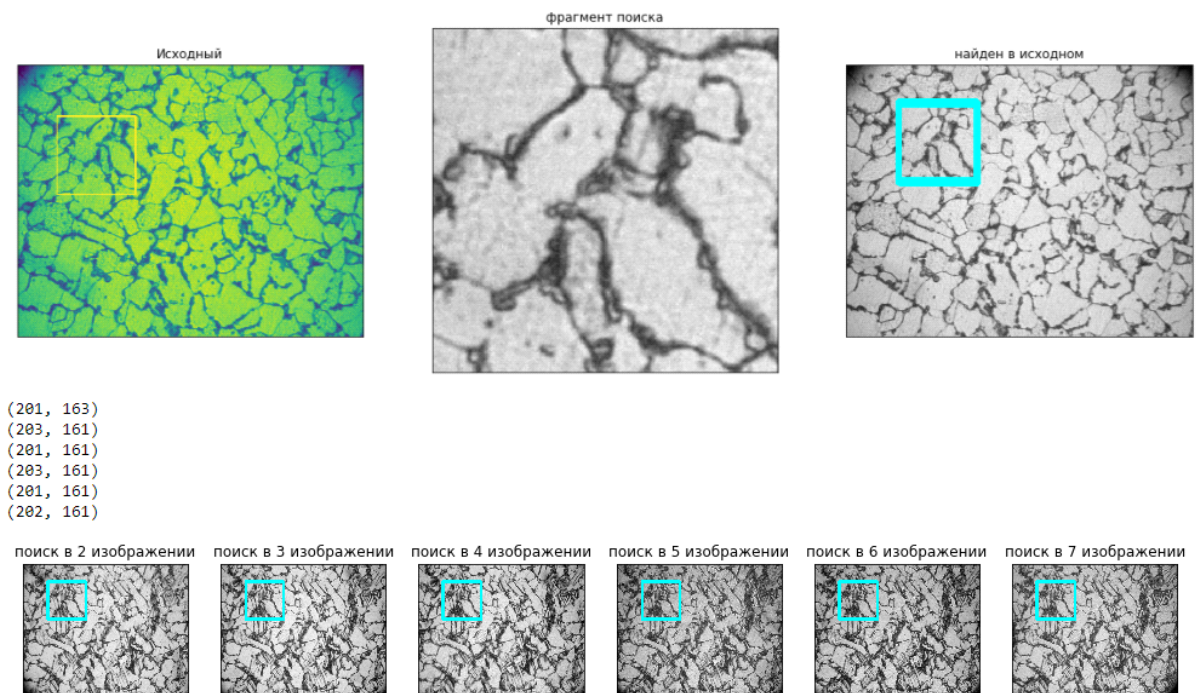


Рис. 5. Поиск фрагментов.

На выходе получаем вероятности нахождения шаблона в изображении. Для получения наиболее лучшего совпадения следует воспользоваться функцией `cv2.minMaxLoc()`.

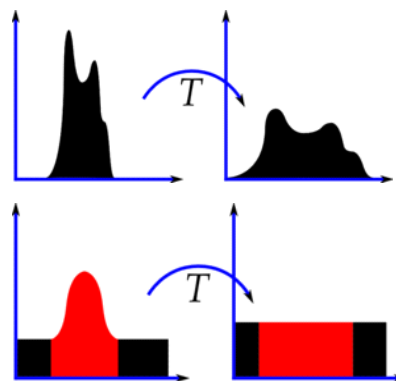


Рис. 6. Выравнивание гистограммы

Для увеличения глобального контраста на многих изображениях используется метод выравнивания гистограмм, особенно эффективен метод, когда используемые данные изображения представлены близкими значениями контраста. Благодаря этой настройке можно лучше распределить интенсивности на гистограмме. Это позволяет областям с более низким локальным контрастом получить более высокий контраст. Выравнивание гистограммы позволяет эффективно распределять наиболее частые значения интенсивности.

Этот метод может увеличить контраст фонового шума, уменьшая полезный сигнал. Такой способ обработки информации полезен для изображений с ярким или темным фоном и передним планом. Преимуществом является то, что расчет изображения не требует больших вычислительных ресурсов. Недостаток метода в том, что он неизбирательный.

Серьезной проблемой может стать побочный эффект, при котором алгоритм дает нереалистичные эффекты на фотографиях. Также выравнивание гистограммы может вызывать нежелательные эффекты (например, видимый градиент изображения), при применении к изображениям с низкой глубиной цвета. Например, если применить его к 8-

битному изображению, отображаемому с 8-битной палитрой серого, это еще больше уменьшит глубину цвета (количество уникальных оттенков серого) изображения. Выравнивание гистограммы лучше всего работает при применении к изображениям с гораздо большей глубиной цвета чем размер палитры, например, непрерывные данные или 16-битные полутонные изображения.

Есть два способа продумать и реализовать выравнивание гистограммы: либо как изменение изображения, либо как изменение палитры. В большинстве случаев лучше изменить палитру, так как при этом сохраняются исходные данные.

Чаще всего можно получить следующий результат, имея входное изображение, показанное на рис. 7.

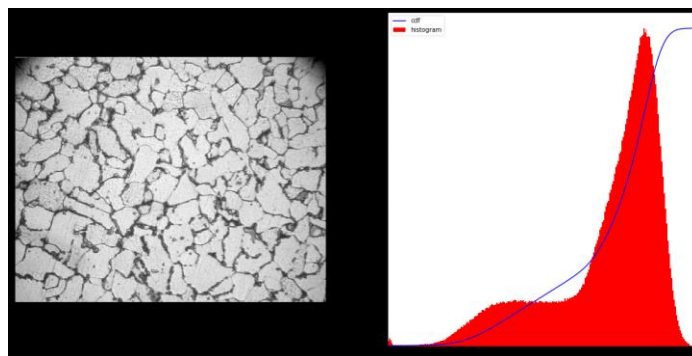


Рис. 7. Исходное изображение.

График для пикселей из фотографии можно увидеть на рисунке 7, на котором синяя линия — это сумма пикселей накопительной суммой, красный график — это количество пикселей по яркости цвета.

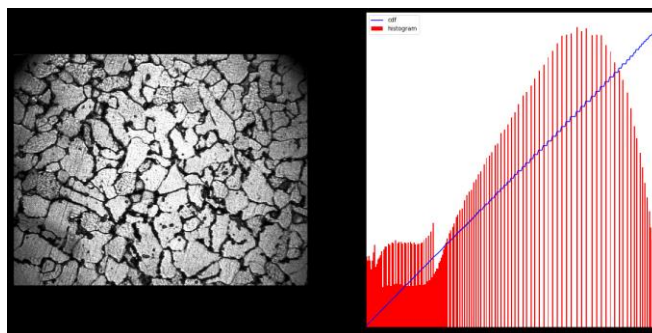


Рис. 8. Результат работы алгоритма

Получившееся изображение выглядит явно контрастнее и с меньшими локальными заветами и затемнениями.

Более продвинутый алгоритм обработки изображения – это CLAHE. CLAHE - адаптивное выравнивание гистограммы. В этом случае изображение разделено на небольшие блоки, называемые «плитками» (размер плитки по умолчанию в OpenCV равен 8x8). Затем каждый из этих блоков, как обычно, выравнивается гистограмма. Результат выполнения алгоритма показан на рисунке 9, рисунок 10 покажет накопленную статистику по пикселям. Проанализировав графики видно, что наиболее ровный график кумулятивной суммы пикселей представляет алгоритм CLACHE.

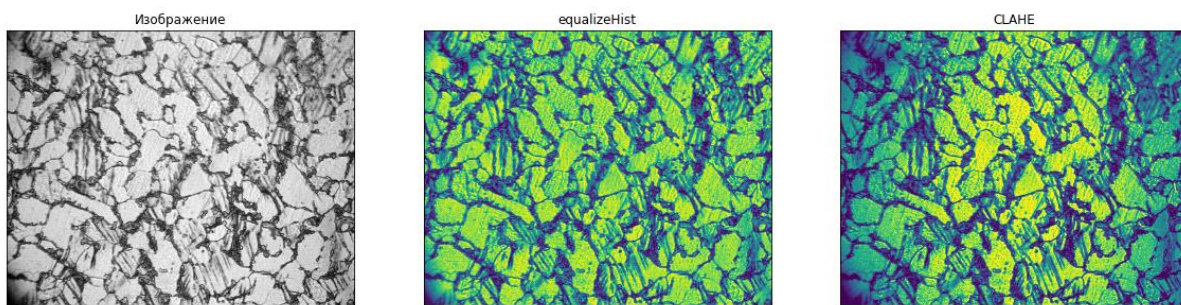


Рис. 9. Три типа изображений

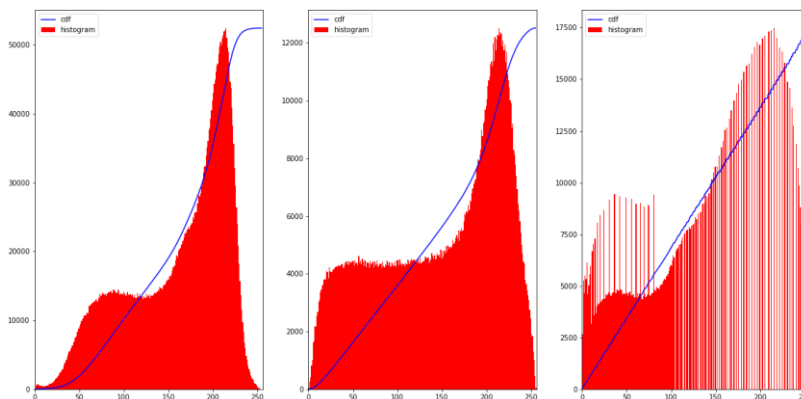


Рис. 10. Графики изображений

Заключение

В работе были рассмотрены разные способы и алгоритмы для увеличения качества изображения и уменьшения внешнего воздействия на получившийся результат. Были приведены результаты работы алгоритма CLAHE и equalizeHist. Изображения из датасета были выровнены по яркости и изменена контрастность. Приведен алгоритм борьбы с изменением положения микроскопа над тестируемым изображением.

Библиографический список

1. Гончар А.В. Исследование микропластической деформации конструкционной стали на начальном этапе усталостного нагружения методами неразрушающего контроля [Текст] /А.В. Гончар, А.Л. Руденко, В.В. Мишакин// Деформация и разрушение материалов. -2012. - V.7. б.м.: - С.37-42.
2. Андреева О.В. Модель накопления поврежденности на изображениях микроструктуры поверхности металлов и сплавов на базе клеточных автоматов [Текст] /О.В. Андреева // Системы управления и информационные технологии, №4.1(66), 2016. – С. 114-117.

A. S. Tsilikh, O.V. Andreeva

TRANSFORMING AND ENHANCING IMAGES FOR NEURAL NETWORKS

Nizhny Novgorod state technical university R.E. Alekseeva

This article examined algorithms for the analysis, operation and transformation of images for subsequent use in neural networks. A brightness and contrast equalization algorithm was applied to process the images and get the best result. An algorithm for compensating for changes in the position of the microscope during control photographs is shown.

Keywords: image transformation, contrast enhancement, fragmentary search, preprocessing for neural networks.

ИССЛЕДОВАНИЕ АЛГОРИТМОВ ДЛЯ РЕШЕНИЯ ЗАДАЧИ ВИДЕОДЕТЕКТИРОВАНИЯ УЧАСТНИКОВ ДОРОЖНОГО ДВИЖЕНИЯ

Нижегородский государственный технический университет им. Р.Е. Алексеева

В данной работе рассмотрены алгоритмы нейронных сетей, чаще всего используемые для решения задачи детектирования объектов. Проведен сравнительный анализ точности и скорости работы выбранных алгоритмов. Определены алгоритмы, являющиеся наиболее выгодными для решения проблемы нахождения и отслеживания транспортного средства в видеопотоке.

Ключевые слова: нейронная сеть, глубокое обучение, детектирование объектов, трекинг.

Введение

Задача детектирования объектов последнее время имеет широкую популярность. Это обосновано ее актуальностью применения в различных сферах: от детектирования личности, до нахождения транспортного средства на дороге. Обнаружение транспортных средств на изображениях является необходимым элементом систем помощи водителю, мониторинга дорожной обстановки, беспилотного управления автомобилями. Решение этой задачи может сделать возможным информирование водителя о текущей обстановке на дороге, например, о приближающихся автомобилях или других участниках дорожного движения. Транспортные средства имеют большое разнообразие типов, отличающихся цветом, формой, поэтому на практике задача их обнаружения и классификации до сих пор остается открытой.

За последние годы достигнут значительный прогресс в области обнаружения объектов с использованием сверточных нейронных сетей. Современные детекторы на основе этих сетей, такие как R-FCN, Faster R-CNN, Multibox, SSD и YOLO стали достаточно быстрыми для использования в потребительских продуктах и для работы на мобильных устройствах.

Цель работы состоит в анализе существующих методов детектирования объектов, и выбор наилучшего решения, по показателям точности и скорости, чтобы найти подходящий вариант для задачи отслеживания транспортного средства в видеопотоке.

Подходы к решению задачи

Проблему обнаружение объектов на изображении можно рассматривать как задачу классификации или как задачу регрессии. В подходе на основе классификации изображение делится на участки, содержащие объект, отличный от фона. Каждый участок пропускается через классификатор, как самостоятельное изображение, чтобы определить класс объекта, изображенного на данном участке. В подходе, использующем регрессию, все изображение пропускается через сверточную нейронную сеть, чтобы произвести обрамляющие окна для объектов на изображении. Рассмотрим алгоритмы, использующие эти подходы.

Алгоритмы на основе предположений регионов. Алгоритмы Fast R-CNN, Faster R-CNN основаны на подходе region proposals, т.е. предположений регионов. Они состоят из двух частей, первая часть генерирует наборы участков, на которых вероятно изображен объект, отличный от фона. Вторая часть обрабатывает эти предположения и классифицирует объект на каждом участке.

Faster R-CNN считается одним из самых точных нейросетевых алгоритмов обнаружения объектов на изображении, но для достижения высокой точности приходится жертвовать скоростью обработки, что не подходит для использования в онлайн-режиме.

Алгоритмы, обрабатывающие изображение целиком. Представителем другого подхода к решению задачи обнаружения объектов является алгоритм YOLO. В этом подходе нейронная сеть предсказывает обрамляющие окна и вероятности классов, применяясь к полному изображению. YOLO является одним из быстрых алгоритмов, но имеет низкую точность на больших и маленьких объектах. [3]

Новейший метод SSD. Он был создан на базе Faster R-CNN и YOLO, однако учел их недостатки, благодаря чему смог добиться новых вершин точности и скорости. Метод основан на сверточной сети прямого распространения, которая создает конечный набор ограничивающих прямоугольных рамок и количественные оценки присутствия в этих рамках объектов различных классов, после чего производится подавление немаксимумов для получения окончательных предсказаний.

Алгоритмы, основанные на предположении регионов

Рассмотрим понятие сверточной нейронной сети и нейросетевые алгоритмы на основе предположений регионов, которые прошли путь от алгоритма R-CNN к алгоритму Faster R-CNN.

Сверточная нейронная сеть. Отличие сверточной нейронной сети от обычной в том, что CNN содержит слои со сверточной архитектурой. В сверточном слое каждый узел соединен только с ограниченной областью предыдущего слоя - полем восприимчивости. В случае обработки изображений, изображение представляет собой набор векторов размера три (R-red G-green B-blue) для каждого пикселя, поэтому устройство CNN таково, что каждый узел соединен с полем восприимчивости - квадратом размера $n \times n$ непосредственно на значениях пикселей. При этом эти поля смещаются с некоторым фиксированным для всех узлов шагом и могут пересекаться.

R-CNN. Система обнаружения объектов R-CNN работает в три этапа. Первый этап генерирует порядка 2000 областей, которые с высокой долей вероятности содержат какой-либо объект, отличный от фона - region proposals. В R-CNN на этом этапе используется алгоритм Selective Search или взаимозаменяемые. Каждое из окон перемасштабируется и обрезается в изображение размера 227×227 пикселей и подается на вход CNN, которая в свою очередь пропускает его через пять сверточных слоев, два полносвязных слоя и выдает в виде вектора признаков размерности 4096. Финальный этап представляет собой набор линейных SVM (support vector machine), которые причисляют каждый вектор какому-либо классу и отдельный регрессор обрамляющих окон.

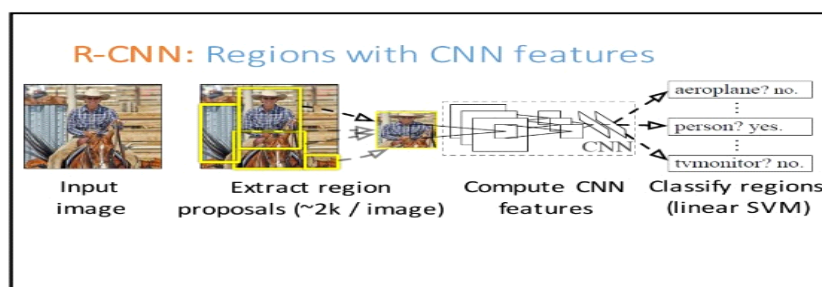


Рис. 1. Схема R-CNN

R-CNN имеет ряд недостатков, самый серьезный заключается в том, что требуется порядка двух-трех дней работы GPU, чтобы обучить алгоритм. При обработке требуется порядка 47 секунд на изображение, что тоже очень долго. Таким образом, основное слабое место заключается в том, что каждое отдельное окно-кандидата требуется пропустить через CNN.

Fast R-CNN. Fast R-CNN предложен в 2015 году с улучшенным по сравнению с алгоритмом R-CNN временем обучения и временем обработки, а также с возросшей точностью обнаружения mAP с 62% до 66% на VOC 2012. Скорость обучения была увеличена в 9 раз, а скорость обработки в 213 раз, чем в оригинальной реализации R-CNN. [1]

Нейросеть получает на вход изображение и набор окон-кандидатов. Несколько сверточных слоев и слоев max pooling производят карту признаков (feature maps). Следующий слой RoI (region of interest) выделяет один вектор признаков из этой карты для каждого предположения объекта. Этот вектор признаков пропускается через серию полносвязных слоев и передается в два выходных слоя: слой softmax, производящий оценку вероятности по каждому классу; слой регрессора обрамляющих окон, который возвращает уточненные позиции обрамляющего окна.

Fast R-CNN показывает хорошие результаты для скорости обработки изображения, но он все еще опирается на медленные алгоритмы, генерирующие окна кандидаты. Selective Search обычно требует порядка 1-2 секунд на изображение, тем самым ограничивая частоту обработки значением в 0.5 FPS, тем самым сводя на нет все дальнейшие улучшения скорости работы сверточной нейросети.

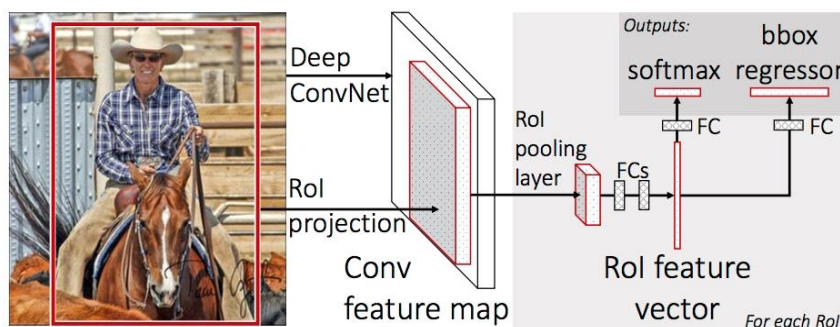


Рис. 2. Схема Fast R-CNN

	R-CNN	Fast R-CNN
Training Time:	84 hours	9.5 hours
(Speedup)	1x	8.8x
Test time per image	47 seconds	0.32 seconds
(Speedup)	1x	146x
mAP (VOC 2007)	66	66.9
Test time per image with Selective Search	50 seconds	2 seconds
(Speedup)	1x	25x

Рис. 3. Результаты Fast R-CNN в сравнении с R-CNN

Faster R-CNN. Алгоритм Faster R-CNN является логическим продолжением алгоритма Fast R-CNN, в котором решены проблемы долгого и не очень точного поиска окон-кандидатов. Тем самым, кроме значительного ускорения работы увеличилась

средняя точность mAP с 66% до 70,4% на VOC 2012. Faster R-CNN использует для первого этапа собственную глубокую полносвязную нейросеть RPN (Region Proposal Networks), а для второго алгоритм Fast R-CNN. [2]

Идея RPN основана на понимании того, что Fast R-CNN требует предположения по окнам-кандидатам только после первых сверточных слоев, а карты признаков, которые получаются после этих первых слоев, могут быть использованы для генерации окон-кандидатов. Большое количество вычислений может быть сокращено, если внести этап поиска регионов в нейросеть. RPN строится путем добавления двух новых сверточных слоев поверх общих сверточных уровней Fast R-CNN. Так как большая часть вычислений в RPN используется совместно с сетью обнаружения, только стоимость двух дополнительных уровней влияет на скорость всего алгоритма Faster R-CNN.

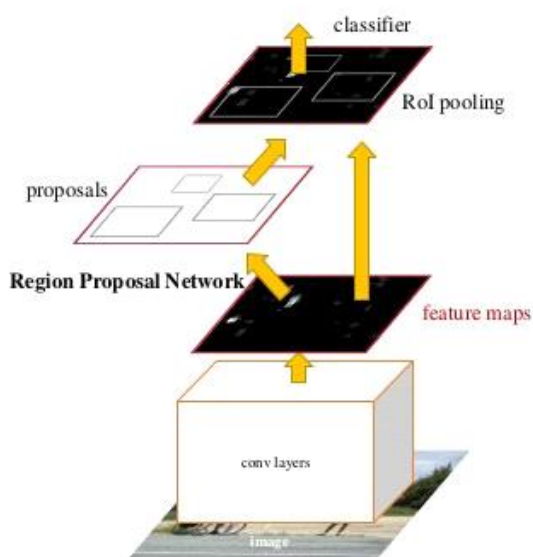


Рис. 4. Схема Faster R-CNN

	R-CNN	Fast R-CNN	Faster R-CNN
Test time per image (with proposals)	50 seconds	2 seconds	0.2 seconds
(Speedup)	1x	25x	250x
mAP (VOC 2007)	66	66.9	66.9

Рис. 5. Сравнительные характеристики

Метод SSD

Структура сети показана на рисунке ниже и может быть разделена на четыре функциональных части. Вход сети (блок 1) - принимает трехканальное (цветное) изображение размером 300x300. Базовая сеть (блоки 2-4) - часть, основанная на стандартной архитектуре, используемой для классификации изображений, в данном случае это VGG-16. Блок 2 полностью повторяет соответствующую часть стека сверток VGG-16, блоки 3 и 4, являясь сверточными слоями, выполняют функции полносвязных слоёв VGG-16, не считая выходного. Слои дополнительных признаков (блоки 4-8) - это сверточные слои и слой подвыборки, которые образуют карты признаков различного разрешения для обнаружения объектов различного масштаба. Далее выполняются операции слияния выходов различных слоёв и подавления немаксимумов для получения выхода сети (блоки 9, 10).

Выходы со слоев блоков 2 и 4-8, являющиеся картами признаков и подвергающиеся слиянию, показаны стрелками, входящими в блок 9. Перед выполнением

принадлежности части изображения, ограниченной рамкой, каждому из классов, и 4 численных значения, определяющих изменение положения и размеров рамки. Затем производится подавление немаксимумов, после чего формируется выход — список значений (класс, степень достоверности принадлежности объекта к классу, x_{min} , y_{min} , x_{max} , y_{max}), где (x_{min}, y_{min}) — координаты левого верхнего угла рамки, (x_{max}, y_{max}) — координаты правого нижнего угла рамки.

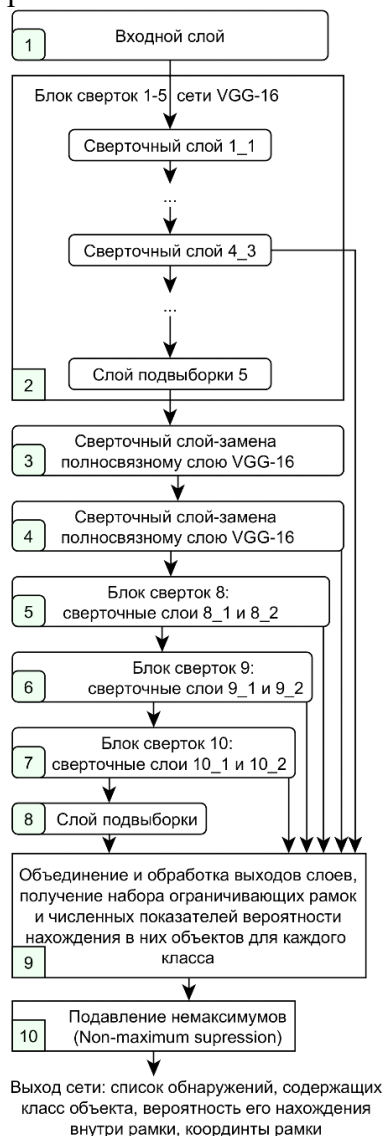


Рис. 6. Архитектура глубокой сверточной нейронной сети SSD

Метод SSD обладает следующими ключевыми особенностями:

Обнаружение объектов различного масштаба. К базовой сети последовательно добавляется несколько сверточных слоев различного размера, которые позволяют получить карты признаков различного разрешения и, соответственно, прогнозировать объекты в различном масштабе (в отличие от, например, YOLO, которая получает лишь одну карту признаков).

Сверточные предсказатели для обнаружения. Каждый слой признаков порождает конечный набор предположений, касающихся класса и положения объекта, используя набор сверточных фильтров. Каждый такой сверточный фильтр имеет ядро, определяющее вероятности принадлежности к классам и смещения рамки по умолчанию.

Ограничивающие рамки по умолчанию. Набор рамок по умолчанию связан с картами признаков различного разрешения, описанных в пункте 1: каждой ячейке карты признаков, соответствует рамка из набора по умолчанию. Для каждой ячейки карты

признаков определяются оценки принадлежности к классам объектов, а также 4 смещения рамки по умолчанию относительно её исходного положения. Объединение этих свойств позволило эффективно дискретизировать множество различных форм результирующих ограничивающих рамок, что положительно сказалось на скорости работы сети.

Method	mAP	FPS	batch size	# Boxes	Input resolution
Faster R-CNN (VGG16)	73.2	7	1	~ 6000	~ 1000 x 600
Fast YOLO	52.7	155	1	98	448 x 448
YOLO (VGG16)	66.4	21	1	98	448 x 448
SSD300	74.3	46	1	8732	300 x 300
SSD512	76.8	19	1	24564	512 x 512
SSD300	74.3	59	8	8732	300 x 300
SSD512	76.8	22	8	24564	512 x 512

Рис. 7. Результаты теста Pascal VOC2007. SSD300 – единственный метод обнаружения в реальном времени, который может достигать более 70% MAP. Благодаря использованию входного изображения большего размера SSD512 превосходит все методы по точности, сохраняя при этом скорость, близкую к скорости реального времени.

Структура MobileNet

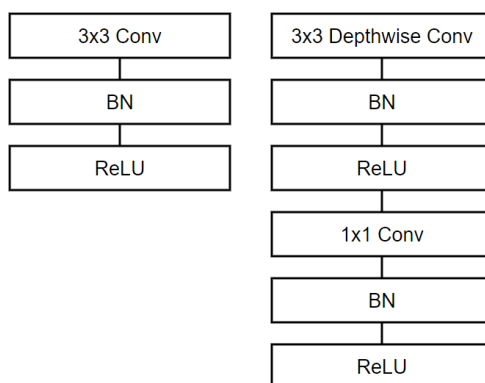


Рис. 8. Слева нарисован блок обычной сверточной сети, а справа — базовый блок MobileNet.

Сверточная часть интересующей нас сети состоит из одного обычного свёрточного слоя с 3x3 свёрткой в начале и тринадцати блоков, изображенных справа на рисунке, с постепенно увеличивающимся числом фильтров и понижающейся пространственной размерностью тензора. Особенностью данной архитектуры является отсутствие max pooling-слоёв. Вместо них для снижения пространственной размерности используется свёртка с параметром stride, равным 2. Двумя гиперпараметрами архитектуры MobileNet являются α (множитель ширины) и ρ (множитель глубины или множитель разрешения). Множитель ширины отвечает за количество каналов в каждом слое. Например, $\alpha=1$ даёт нам архитектуру, описанную в статье, а $\alpha=0.25$ — архитектуру с уменьшенным в четыре раза числом каналов на выходе каждого блока.

Множитель разрешения отвечает за пространственные размеры входных тензоров. Например, $\rho=0.5$ означает, что высота и ширина feature map, подаваемой на вход каждому слою будет уменьшена вдвое. Оба параметра позволяют варьировать размеры сети: уменьшая α и ρ , мы снижаем точность распознавания, но в то же время увеличиваем скорость работы и уменьшаем потребляемую память. [4]

MobileNetV2

Появление MobileNet уже само по себе сделало революцию в компьютерном зрении на мобильных платформах, но MobileNetV2 — следующее поколение нейросетей этого семейства, которое позволяет достигать примерно такой же точности распознавания при ещё большей скорости работы. Основной строительный блок этой сети в целом

похож на предыдущее поколение, но имеет ряд ключевых особенностей. Как и в MobileNetV1, здесь есть сверточные блоки с шагом 1 (на рисунке слева) и с шагом 2 (на рисунке справа). Блоки с шагом 2 предназначены для снижения пространственной размерности тензора и, в отличие от блока с шагом 1, не имеют residual connections.

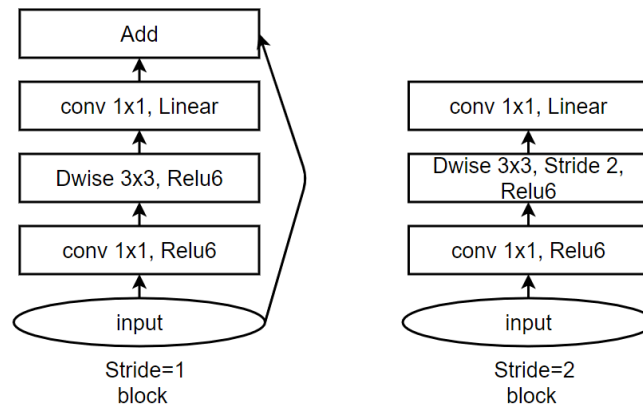


Рис. 9. Сверточные блоки MobileNetV2

Блок MobileNet, называемый расширяющим сверточным блоком, состоит из трёх слоёв. Сначала идёт pointwise convolution с большим количеством каналов, называемый expansion layer. На входе этот слой принимает тензор размерности $D_f * D_f * C_{in}$, а на выходе выдает тензор $D_f * D_f * (t * C_{in})$, где t — новый гиперпараметр, названный уровнем расширения. Этот слой создает отображение входного тензора в пространстве большей размерности. Затем идёт depthwise convolution с ReLU6-активацией. Этот слой вместе с предыдущим образует уже знакомый строительный блок MobileNetV1. На входе этот слой принимает тензор размерности $D_f * D_f * (t * C_{in})$, а на выходе выдает тензор $(D_f/s) * (D_f/s) * (t * C_{in})$, где s — шаг свертки, depthwise convolution не меняет число каналов. В конце идёт 1×1 -свертка с линейной функцией активации, понижающая число каналов. На входе такой слой принимает тензор размерности $(D_f/s) * (D_f/s) * (t * C_{in})$, а на выходе выдает тензор $(D_f/s) * (D_f/s) * C_{out}$, где C_{out} — количество каналов на выходе блока. Фактически, именно третий слой в этом блоке, называемый bottleneck layer, и является основным отличием второго поколения MobileNet от первого.

SSDLite-архитектура для задачи object detection, использующая MobileNetV2 в свёрточной части, превосходит известный детектор реального времени YOLOv2 по точности на датасете MS COCO, при этом показывая в 20 раз большую скорость и в 10 раз меньший размер. Если объединить архитектуру MobileNet и структуру Single Shot Detector (SSD), получится быстрый и эффективный метод обнаружения объектов на основе глубокого обучения.

В итоге используется именно MobileNet в SSD, потому что Resnet, VGG или alexnet имеет большой размер сети и увеличивает количество вычислений, тогда как в MobileNet существует простая архитектура, состоящая из глубокой свертки 3×3 , за которой следует точечная свертка 1×1 .

Model name	Speed (ms)	COCO mAP[¹]
ssd_mobilenet_v1_coco	30	21
ssd_mobilenet_v1_0.75_depth_coco	26	18
ssd_mobilenet_v1_quantized_coco	29	18
ssd_mobilenet_v1_0.75_depth_quantized_coco	29	16
ssd_mobilenet_v1_ppn_coco	26	20
ssd_mobilenet_v1_fpn_coco	56	32
ssd_resnet_50_fpn_coco	76	35
ssd_mobilenet_v2_coco	31	22
ssd_mobilenet_v2_quantized_coco	29	22
ssdlite_mobilenet_v2_coco	27	22
ssd_inception_v2_coco	42	24
faster_rcnn_inception_v2_coco	58	28
faster_rcnn_resnet50_coco	89	30
rfcn_resnet101_coco	92	30
faster_rcnn_resnet101_coco	106	32

Рис. 10. Оценка скорости модели

На рис. 10 указано время работы в мс для каждого изображения 600x600 (включая всю предварительную и постобработку), но эти тайминги сильно зависят от конкретной конфигурации оборудования (эти тайминги были выполнены с использованием Nvidia GeForce GTX TITAN X).

Заключение

Были рассмотрены алгоритмы Fast R-CNN, Faster R-CNN, YOLO и SSD для детектирования изображений. Аналитическим путем установлено, что для задачи нахождения объекта в реальном времени наиболее приемлемым вариантом по показателю времени и качества является SSD. Изучена модель SSD с MobileNet, которая показывает лучшие данные на датасете COCO.

Библиографический список

1. Girshick R. Fast R-CNN – Текст: электронный // URL: <https://arxiv.org/pdf/1504.08083.pdf> (дата обращения: 12.12.2020).
2. Ren S., He K., Girshick R., Sun J. Faster R-CNN: Towards Real-Time Object Detection with Region Proposal Networks – Текст: электронный // URL: <https://arxiv.org/pdf/1506.01497.pdf> (дата обращения: 18.12.2020).
3. Redmon J., Divvala S., Girshick R., Farhadi A. You Only Look Once: Unified, Real-Time Object Detection – Текст: электронный // URL: <https://arxiv.org/pdf/1506.02640v5.pdf> (дата обращения: 20.12.2020).
4. Panchenko D. MobileNet: меньше, быстрее, точнее – Текст: электронный // URL: <https://habr.com/ru/post/352804/> (дата обращения: 22.12.2020).

E.A. Totmenina, O.V. Andreeva

**INVESTIGATION OF ALGORITHMS FOR SOLVING THE TASK OF VIDEO
DETECTION OF ROAD USERS**

Nizhny Novgorod state technical university R.E. Alekseeva

In this paper, the algorithms of neural networks, most often used to solve the problem of object detection, were considered. A comparative analysis of the accuracy and speed of the selected algorithms is carried out. The algorithms that are the most advantageous for solving the problem of finding and tracking a vehicle in a video stream have been determined.

Keywords: neural network, deep learning, object detection, tracking.

ОБРАБОТКА БОЛЬШИХ ДАННЫХ С ЦЕЛЬЮ СОЗДАНИЯ СИСТЕМЫ ОПРЕДЕЛЕНИЯ ГЕОПОЗИЦИИ ПО ФОТОГРАФИЯМ МЕСТНОСТИ

Нижегородский государственный технический университет им. Р.Е. Алексеева

В работе рассматривается задача обработки больших данных. Данные относятся к изображениям улиц. Описывается процесс разработки алгоритма для определения геопозиции с использованием изображения местности на основе алгоритмов глубокого обучения и кластеризации данных по геопозиции.

Ключевые слова: большие данные, поиск геопозиции по фотографии, задача кластеризации, нейронная сеть векторизации данных.

Введение

Определение географического положения по фотографии местности является очень сложной задачей, поскольку изображения, сделанные в разных уголках мира, имеют огромное количество вариаций даже для одной местности: например, фотографии сделаны в разное время суток, на них изображены разные объекты или при фотографировании использованы разные настройки камеры. Кроме того, изображения местности часто неоднозначны и поэтому дают очень мало визуальных подсказок об их реальном местоположении. В связи с этим задачу геолокализации фотографий часто упрощают, ограничивая проблему городскими фотографиями, например, фотографиями хорошо известных достопримечательностей. Из-за сложности задачи и несбалансированного распределения фотографий, снятых в разных точках земного шара, методы, основанные на использовании сверточных нейронных сетей (CNNs), рассматривают фотогеолокализацию как классификационную задачу, разделяющую поверхность земли на географические ячейки с одинаковым количеством изображений.

Целью представленной работы является разработка алгоритма позиционирования человека в пространстве по сделанным им фотографиям местности на основе методов компьютерного зрения и сравнение полученных результатов с существующими системами позиционирования человека.

Алгоритм позиционирования человека

В соответствии с разработанным алгоритмом для получения наиболее точной геолокации человека по фотографиям местности, в которой он находится, необходимо проделать несколько шагов (рис.1). Во-первых, получить начальные данные - в нашем случае это фотография местности и начальные координаты, которые будут определены с помощью датчика GPS. Далее необходимо преобразовать полученные данные: для этого будут произведены операции для нормализации фотографий. На следующем этапе изображение будет отправлено на переобученную свёрточную нейронную сеть, построенную на архитектуре ResNet [1], с целью выделения определяющих признаков на фотографии, таких как: дорожные знаки, здания, достопримечательности и т.п. Далее полученные данные с помощью метода k-ближайших соседей (k-means) будут сравниваться с базой данных улиц и городов с целью определения похожей улицы. На последнем этапе происходит корректировка местоположения человека.

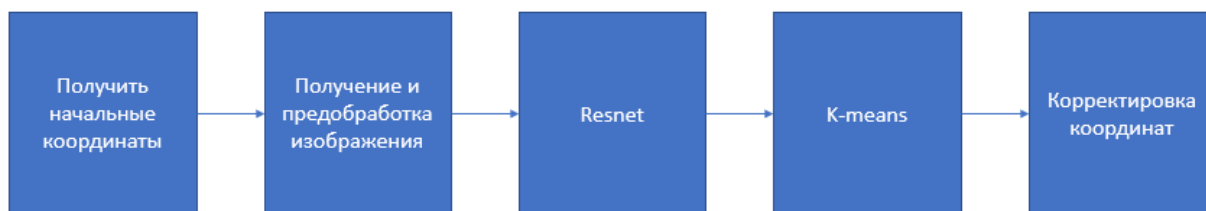


Рис. 1. Этапы работы алгоритма позиционирования человека

Предложенный подход предполагает, что изображения местности находятся в географической близости друг от друга и имеют общие элементы (дорожные знаки, номера домов, архитектурные строения), что делает их визуально похожими. Поэтому предлагается сначала получить несколько кандидатов (фотографий района) из базы фотографий, а затем оценить каждого кандидата на основе визуального сходства, имеющегося на нем изображения с запросом (фотографией пользователя). Если визуальное сходство высокое, это может служить дополнительным доказательством того, что кандидат визуально похож на запрос.

Набор данных

С целью тестирования пригодности алгоритма для задачи геолокации было принято решение проводить исследования на ограниченной местности, - двух районах Нижнего Новгорода. Для обучения сверточной нейронной сети было собрано 1000 фотографий местности из Googlestreet View [2], содержащих изображения популярных мест в отобранных для исследования районах города, полученных с разных ракурсов и в разные времена года; кроме того, фотографии содержали широту и долготу местности, на которых были сделаны, и угол поворота камеры. Все изображения были поделены на две группы, ассоциирующие два района города.

Применение сверточной нейронной сети

Задача геолокации изображений местности в работе ставится как задача классификации. В связи с этим необходимо поделить городскую местность на множество географических ячеек, что позволит ассоциировать каждую фотографию с определенным районом города; поиск в таких ячейках даст большую вероятность найти похожего кандидата. Входными данными для сверточной сети ResNet является изображение, которое содержит отличительные характеристики местности, а целевым выходом – одномерный вектор, кодирующий ячейку, содержащую гео-метку изображения. Учитывая информацию на полученном изображении местности, выход этой модели представляет собой распределение вероятностей соответствия определенному району по всему городу. Преимущество этой формулировки перед регрессией от пикселей к координатам широты/долготы состоит в том, что модель может выразить свою неопределенность относительно изображения, присвоив каждой ячейке уверенность в том, что изображение было сделано именно там. В отличие от этого, регрессионная модель была бы вынуждена точно определить одно место и не имела естественного способа выразить неопределенность относительно своего решения.

Для классификации изображений и извлечения меток изображения применяется модель сверточной сети ResNet, содержащая 152 слоя. Модель была обучена более чем на 10000 обучающих изображений местности, сделанных в различных районах и улицах города. Это хорошо согласуется с нашим подходом, поскольку полученный классификатор уже различает изображения, относящиеся к определенным участкам района. Прогнозируется геолокация пользователя на основе всего набора изображений обучающей выборки, используя максимальную вероятность выходного вектора, путем сложения вероятностей всех классов, которые отнесены к одной и той же координатной категории, а также генерируется соответствующий кластер, который будет характеризовать определенный участок местности.

Предлагается проводить обучение сети с использованием архитектуры ResNet. Веса инициализируются с помощью предварительно обученной модели с возможностью интеграции дополнительных слоев. Чтобы избежать переобучения, обучающая выборка изменяется и дополняется случайным выбором изображений из разных групп местности. Обучающая выборка покрывает не менее 70% изображений всей базы данных. Кроме того, входные изображения случайным образом переворачиваются и затем обрезаются до размера 224×224 пикселя. В работе используется стохастический оптимизатор градиентного спуска с начальной скоростью обучения 0,01; импульсом 0,9 и спадом веса 0,0001; - такие параметры были выбраны в ходе экспериментов и являются наиболее оптимальными для обучения. Скорость обучения экспоненциально снижается в 0,5 раза после каждых пяти эпох обучения. Первоначально сеть обучается на 15 эпохах, размер пакета изображений, подаваемых для обучения на одну эпоху, равен 64 изображениям. Работа сети была проверена на 1000 изображениях. Нейронная сеть реализована на языке программирования Python с использованием библиотеки Tensor Flow [3]. В табл. 1 представлены результаты - вероятность точного определения сетью местности похожей на исходное изображение (Топ-1 – вероятность того, что первое изображение является похожей местностью; Топ-5 – вероятность того, что в пяти похожих результатах есть похожее изображение).

Таблица 1

Точность классификации сцен Топ-1 и Топ-5

Сеть	Топ-1	Топ-5
Обученная сеть ResNet	81,7 %	97,5 %

На выходе сети получается вектор из 1000 элементов, характеризующих различные особенности изображения (дорожные знаки, заборы, дома, архитектурные строения и т.п.). Для определения точного местоположения человека происходит сравнение полученного вектора с массивом фотографий из Google street View методом k-means [4].

Реализация метода k– ближайших соседей

Каждый вектор, полученный на выходе обученной сети, будет определён в свой кластер, в результате все изображения в базе данных будут поделены на кластеры, с которыми будет сравниваться обрабатываемое входное изображение. С добавлением в систему дополнительных районов или городов, количество кластеров будет расти.

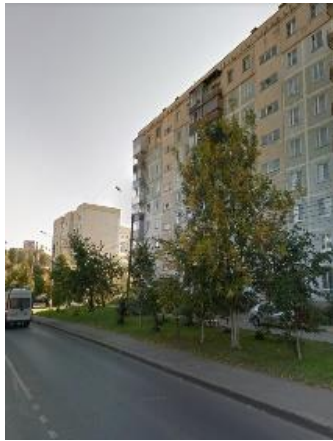
Для реализации классификации местности в работе использовалась технология кластеризации на основе алгоритма k-средних (k-means), в основе которой лежит суммарное квадратичное отклонение центров масс кластера от точек (1):

$$V = \sum_{i=1}^k \sum_{x \in S_i} (x - \mu_i)^2, \quad (1)$$

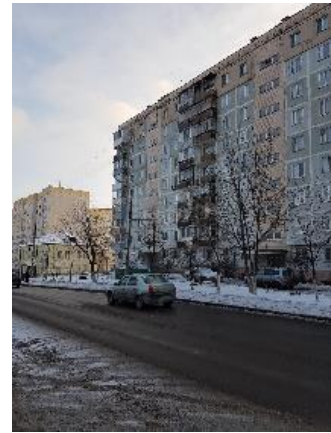
где k – количество кластеров, S_i – i -й кластер, x – точка кластера, μ_i – центр масс i -го кластера.

Кластеры позволяют избежать расчета расстояния от точки, представляющей геотег, поданный на вход системы, до всех остальных точек базы данных в пространстве признаков, и работать лишь в рамках одного кластера.

На рис. 2 и 3 представлены результаты работы обученной нейронной сети: изображения (рис. 2, б; 3, б) сделаны пользователем с фотокамеры, изображения (рис. 2, а; 3, а) найдены с помощью разработанного алгоритма в базе данных и являются наиболее похожими на пользовательские изображения местности.



а)



б)

**Рис. 2. Пример 1 работы,
а) фото из базы данных изображений местности; б) фото с камеры пользователя**



а)



б)

**Рис. 3. Пример 2 работы,
а) фото из базы данных изображений местности; б) фото с камеры пользователя**

Анализ и сравнение полученных результатов

После получения данных от нейронной сети, производится корректировка геолокации пользователя на основе полученных результатов путем сдвига относительно угла поворота камеры, полученная таким образом геолокация – это и есть местоположение фотографии пользователя (или самого пользователя).

В связи с тем, что в качестве эксперимента был взят небольшой участок города (два района), точность определения геопозиции человека с помощью разработанного алгоритма весьма высока. При расширении области поиска, точность может ухудшиться.

Сравним работу полученной системы с уже имеющимися системами определения местоположения человека. На данный момент существует несколько популярных способов определения геопозиции, однако, базовыми, несомненно, являются системы глобального позиционирования GPS и системы сотовой связи, определяющие местоположение телефона. В табл. 2 представлен сравнительный анализ существующих технологий определения геопозиции с результатами, полученными разработанным методом.

Сравнительный анализ

Тип	Точность, м
GPS	6 – 15
Сотовая связь	100 – 500
Разработанный алгоритм	0,5 - 2

Заключение

В работе предложен новый подход к определению местоположения человека, основанный на поиске локации по изображению местности, который состоит из трех этапов, а именно: получение изображения и предобработка, построение гео-окрестностей, ранжирование изображений. Были проведены эксперименты для сравнения эффективности предложенного подхода с двумя базовыми способами определения геолокации.

В результате работы была создана система, реализованная на языке программирования Python 3, с использованием фреймворка Tensorflow, позволяющая получить местоположение человека по фотографии местности. В качестве исходной сети использовалась сверточная нейронная сеть ResNet-50. Обучение применялось на собранном наборе данных размером 10000 фото. Поиск местности осуществляется по двум районам города Нижний Новгород. Точность работы сети составила 81%. Разработанная система позволяет получить весьма точное позиционирование человека - с ошибкой менее 1 м. В качестве направлений дальнейшей работы планируется расширение области поиска и улучшение показателей точности.

Библиографический список

1. Resnet: остаточная CNN для классификации изображений // URL: <https://neurohive.io/ru/vidy-nejrosetej/resnet-34-50-101/> (дата обращения 02.01.2021)
2. GooglestreetviewAPI [Электронный ресурс]. // URL: <https://developers.google.com/maps/documentation/streetview/overview> (дата обращения 02.01.2021)
3. Tensorflowguide [Электронный ресурс]. // URL: <https://www.tensorflow.org/guide?hl=en> (дата обращения 02.01.2021)
4. K-means clustering in Machine Learning [Электронный ресурс]. // URL: <https://towardsdatascience.com/understanding-k-means-clustering-in-machine-learning-ba6e67336aa1> (дата обращения 02.01.2021)

O.P. Timofeeva, A.A. Dmitriev, M.B. Bagirov

PROCESSING OF BIG DATA IN ORDER TO CREATE A SYSTEM FOR ESTIMATING THE LOCATION OF A PHOTO

Nizhny Novgorod state technical university. R.E. Alekseev

The paper deals with the problem of big data processing. Data refer to the images of the streets. The article describes the process of developing an algorithm for correcting a geo-location using an image of the terrain using deep learning algorithms and clustering data on the geo-location.

Keywords: big data, geo-location search by photo, clustering problem, neural network of data vectorization.

РАЗРАБОТКА И ИССЛЕДОВАНИЕ АЛГОРИТМОВ АВТОМАТИЧЕСКОГО РАСПОЗНОВАНИЯ АВТОМОБИЛЬНЫХ НОМЕРОВ

Нижегородский государственный технический университет им. Р.Е. Алексеева

В работе рассматривается задача распознавания российских автомобильных номерных знаков, в которой детектирование номера производится с помощью разработанного алгоритма, основанного на использовании специальных матричных фильтров для бинарного изображения. Для сравнения результатов работы алгоритма производится решение задачи детектирования известными методами: методом Виолы–Джонса и методом, использующим нейронную сеть YOLO. Проводится анализ полученных решений, делаются выводы о преимуществах, недостатках и области применимости каждого из рассмотренных подходов.

Ключевые слова: автоматическое распознавание автомобильных номерных знаков, детектирование автомобильных номеров, ANPR, метод Виолы–Джонса, YOLO, детектирование на основе контуров.

Введение

С высоким темпом увеличения степени автоматизации процессов во всех областях человеческой деятельности возникает потребность и в автоматизации сбора визуальной информации. Автоматическое распознавание автомобильных номеров может и уже используется во многих задачах, от мониторинга автомобильного потока до организации деятельности предприятий по обслуживанию автомобилей. Автомобильный номер может являться уникальным идентификатором не только самого автомобильного средства, но и его водителя, и наряду с иными признаками использоваться при автоматизации других процессов. Потому так важно иметь возможность распознавать автомобильные номера не только точно, но и быстро. В данной статье будет представлен быстрый алгоритм детектирования автомобильного номера и его сравнение с другими известными подходами.

Постановка задачи

Задача автоматического распознавания автомобильных номерных знаков (Automatic number-plate recognition - ANPR) включает в себя две подзадачи (рис.1):

- Детектирование номерного знака.
- Распознавание текста на номерном знаке:
 - Детектирование символов.
 - Распознавание символов.



Рис. 1. Разбиение задачи ANPR на подзадачи

В данной статье мы коснемся первой подзадачи, а именно задачи детектирования автомобильных номеров. Представим ее решение методом, основанным на выделении контуров объектов на изображении, разработанным в процессе работы над задачей. Также рассмотрим несколько других известных подходов к решению задачи детектирования с целью сравнения разработанного алгоритма с известными по качеству и скорости работы, а также определения области применимости рассмотренных подходов.

Описание алгоритмов детектирования автомобильных номеров

Алгоритм детектирования на основе контуров объектов на изображении

В ходе работы над задачей детектирования автомобильных номеров разработан собственный алгоритм, не требующий специального предварительного обучения, который основан на морфологических алгоритмах, реализуемых с помощью использования на бинарном изображении специальных матричных фильтров. На рис. 2 приведена схема работы алгоритма.

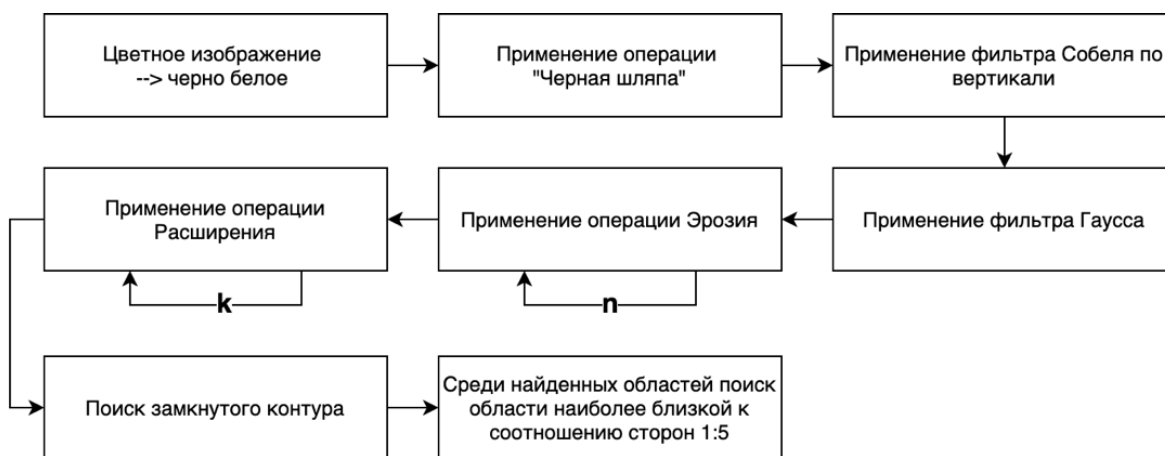


Рис. 2. Схема алгоритма детектирования на основе контуров объектов

Рассмотрим подробнее шаги алгоритма детектирования на основе контуров объектов на изображении:

- Переводим исходное цветное изображение в бинарное изображение.
- Применяем операцию преобразования *«черная шляпа»* [1], которое заключается в использовании операции расширения, а затем операции эрозии, после чего полученное изображение вычитается из исходного.

Операция расширения реализуется с помощью прохождения квадратным фильтром (структурным элементом) по каждому единичному пикселю изображения. При этом нулевые (черные) пиксели изображения, попавшие в область фильтра, меняют свое значение на единицы. Тем самым белая область на бинарном изображении расширяется.

Операция эрозии — это обратная операция, уменьшающая белые области на изображении. При этом белые области меньшего размера, чем фильтр, полностью пропадают с изображения.

Применяя последовательно расширение и эрозию, мы выделим на изображении черные области, находящиеся внутри белых. Если это применить к российскому автомобильному номеру (фон белый, символы черные), используя фильтр с соотношением сторон 1 к 5 (соотношение сторон стандартного российского автомобильного номерного знака), то мы получим выделение символов на номерном знаке (рис.3, а).

- Выделим более явно границы символов номера, используя оператор Собеля для производной вертикали. Применение этого фильтра записывает на месте каждого пикселя вместо интенсивности приближенное значение градиента интенсивности

относительно соседних пикселей. Таким образом пиксели, находящиеся на контуре какого-либо объекта, приобретают наибольшее значение, так как именно на границах интенсивность пикселей резко меняется (рис. 3, б).

- На предыдущих шагах были выделены символы на номерном знаке. Данное изображение можно будет использовать для детектирования каждого символа в задаче распознавания текста на номерном знаке. Однако, сейчас стоит задача в детектировании всего автомобильного номера, поэтому необходимо объединить символы в одну прямоугольную белую область. Для этого сначала смажем изображение, пройдясь по нему фильтром Гаусса (размером 11×11 пикселей) и снова преобразуем изображение к бинарному виду.
- Избавимся от областей на изображении, содержащих малую белую область, для этого два раза проведем операцию эрозии (в общем случае операция проводится n раз). Оставшиеся белые области, наоборот, увеличим, для этого четыре раза применим операцию расширения (в общем случае операция применяется k раз) (рис.3, в).
- На последнем шаге находим контуры всех получившихся белых областей и выбираем ту область, соотношение высоты и ширины которой наиболее близко к российскому автомобильному номеру (1:5) (рис. 3, г).

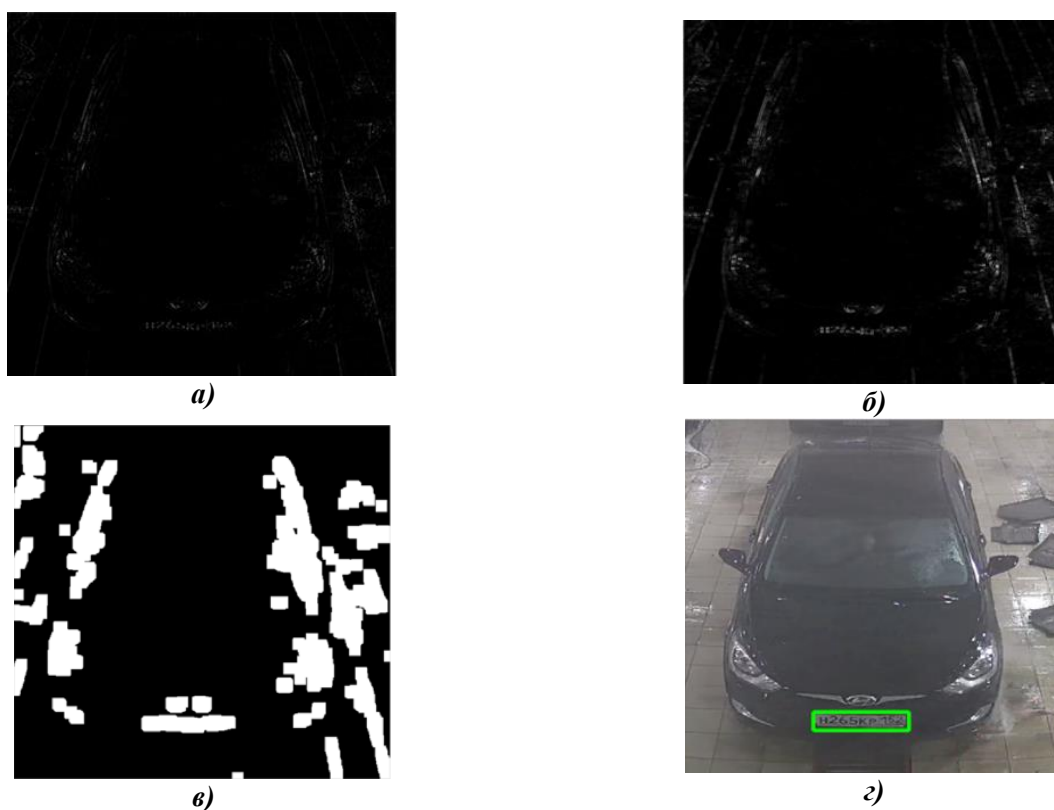


Рис. 3. Пример детектирования на основе контуров объектов на изображении, результаты применения:

- а) преобразования «Черная шляпа»; б) фильтра Собеля;
в) фильтра Гаусса и нескольких операций расширения-эрозии; г) алгоритма в целом

Метод Виолы – Джонса

Рассмотрим известный алгоритм детектирования – метод Виолы-Джонса, в основе которого лежит принцип сканирующего окна и пирамиды изображений. Алгоритм пробегает таким окном по каждому изображению в пирамиде и накладывает на эти области изображения примитивы Хаара, в различных расположениях. Сочетание примитива и его расположения будет являться признаком. Если признаки для рассматриваемого окна совпадают с признаками для автомобильного номера, то алгоритм возвращает координаты данного окна, - это и будет результатом детектирования (рис. 4).

Метод Виолы – Джонса [2] обучается крайне долго и требует высокопроизводительный компьютер, поэтому для проведения сравнительного анализа использовался уже предварительно обученный для российских автомобильных номеров каскад, входящий в состав библиотеки OpenCV.

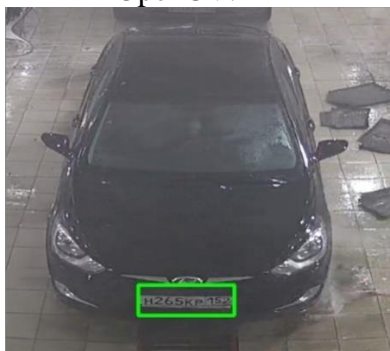


Рис. 4. Пример детектирования методом Виолы – Джонса

Детектирование номеров с помощью сверточной нейронной сети на основе архитектуры YOLOv5

Рассмотрим еще один известный подход - сверточную нейронную сеть YOLO [3], которая за один проход окном по изображению способна детектировать и классифицировать объекты на данном изображении. Для данной работы сеть была обучена на изображениях из свободно распространяемого набора данных «License plate Object Detection», который состоит из фотографий автомобильных номеров размером 416×416 пикселей и текстовых аннотаций к ним. Набор данных был разбит на обучающую и тестовую выборки в соотношении 4 к 1. Обучение происходило на протяжении 100 эпох. Для оптимизации использовался стохастический градиентный спуск.

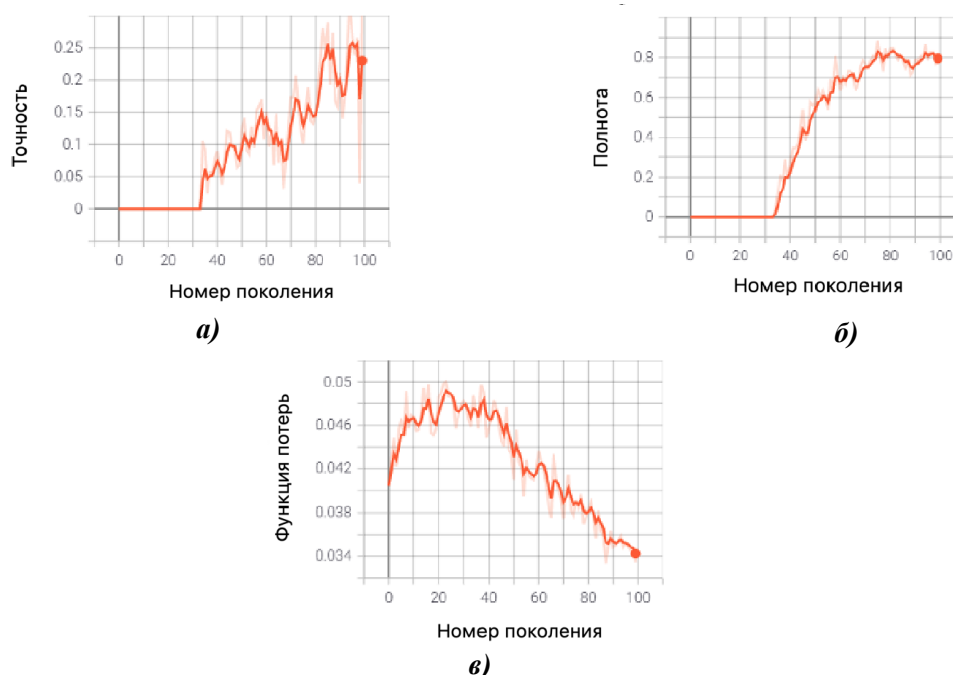


Рис. 5. Динамика изменения в процессе обучения:

а) точности; б) полноты; в) значения функции потерь

На рис. 5 приведены графики изменения точности (рис. 5, а), полноты (на тестовом наборе данных) (рис. 5, б) и значения функции потерь YOLO loss (рис.5, в). На рисунке б приведен пример работы обученной нейронной сети на тестовом изображении.

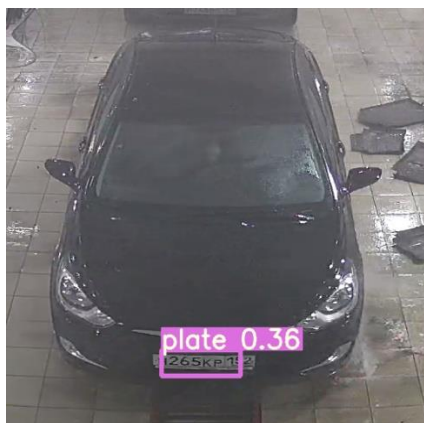


Рис. 6. Пример детектирования автомобильного номера с использованием нейронной сети YOLO

Анализ рассмотренных алгоритмов

Проведем анализ результатов решения задачи детектирования автомобильных номеров представленными в работе методами.

Алгоритм на основе выделения контуров объектов на изображении, разработанный нами, не нуждается в специальном обучении, однако может потребоваться ручной подбор некоторых параметров (например, количество операций эрозии и расширения на шагах 5 и 6), которые будут зависеть от внешних условий освещенности, нормализации интенсивности входного изображения [4] и выставления параметра, определяющего соотношение сторон номерного знака (в зависимости от типа транспортного средства и угла, под которым был сделан кадр). Обобщающая способность у разработанного алгоритма оказалась ниже, чем у двух других рассматриваемых в данной статье, но при этом скорость работы алгоритма наиболее высокая. На основе проведенных тестирований было выявлено, что алгоритм на основе контуров объектов на изображении работает примерно в 5 раз быстрее YOLO, и более чем в 10 раз быстрее метода Виолы – Джонса.

Что касается сравнения двух других методов, то метод Виолы – Джонса требует значительного количества времени на обучение каскадов и сам процесс детектирования занимает в несколько раз больше времени, чем подход с использованием YOLO, так как требует построения и обработки пирамиды изображений. При этом точность детектирования у этих алгоритмов сопоставима. Главным преимуществом использования нейронной сети YOLOv5 является ее высокая обобщающая способность. Этот алгоритм может достаточно точно детектировать автомобильные номера при разных внешних условиях.

Вывод

Таким образом, было рассмотрено и проанализировано три алгоритма детектирования российских автомобильных номерных знаков, один из которых (на основе контуров объектов) был разработан в рамках данной работы. Можно сделать вывод, что если необходимо, чтобы система детектирования корректно работала при разных внешних условиях (например, не в помещении), то лучше всего подходит нейронная сеть на основе архитектуры YOLO, обученная на фотографиях автомобильных номеров. Однако, если необходима высокая скорость работы и есть возможность обеспечить постоянные условия освещенности, то, настроив ряд параметров под конкретные условия, эффективнее будет использовать предложенный в данной работе алгоритм на основе выделения контуров объектов на изображении. В сочетании с быстрым алгоритмом решения задачи детектирования текста на номерном знаке (например, средствами tesseract) это позволит решать задачи ANPR на видеопотоке в рамках близких к реальному времени.

Библиографический список

1. Morphological Transformations // OpenCV URL: https://docs.opencv.org/master/d9/d61/tutorial_py_morphological_ops.html (дата обращения: 19.12.2020)
2. Paul Viola, Michael Jones. Rapid. Object detection using a boosted cascade of simple features. // Accepted conference on Computer Vision and Pattern Recognition. – 2001.
3. YOLOv5 New Version - Improvements and Evaluation // Roboflow URL: <https://blog.roboflow.com/yolov5-improvements-and-evaluation/> (дата обращения: 21.12.2020).
4. Соколова Э.С., Дмитриев Д.В., Корелин О.Н., Кобляков Д.А., Багиров М.Б., Сачков А.Д.//Разработка и исследование алгоритмов калибровки аналоговых камер высокого разрешения для систем помощи водителю//Транспортные системы. 2018. № 3 (9). С. 11-15.

О.Р. Timofeeva, A.N. Raitsin, M.B. Bagirov

DEVELOPMENT AND ANALYSIS OF ALGORITHMS FOR AUTOMATIC LICENSE PLATE RECOGNITION

Nizhny Novgorod State Technical University. R.E. Alekseeva

The paper describes the problem of recognition of Russian car license plates, in which license plate detection is performed using the developed algorithm based on the use of special matrix filters for a binary image. To compare the results of the algorithm, the detection problem also is solved using the methods: The Viola – Jones method and the method using the YOLO neural network. Analyzed and conclusions are drawn about the advantages and disadvantages of each of the considered algorithms.

Keywords: automatic license plate recognition, license plate detection, ANPR, Viola-Jones method, YOLO, edge-based detection

ЦИФРОВАЯ ФИЛЬТРАЦИЯ С ИСПОЛЬЗОВАНИЕМ DSP TI 5535

Нижегородский государственный технический университет им. Р.Е. Алексеева

Цифровая фильтрация является одним из наиболее мощных инструментов средств цифровой обработки сигналов. Кроме очевидных преимуществ цифровые фильтры способны удовлетворять таким техническим требованиям, которые было бы трудно достичь в аналоговом исполнении. В статье рассматриваются способы проектирования и применения цифрового фильтра с помощью цифрового сигнального процессора. Рассчитываются коэффициенты и импортируются в программный код. Оценивается качество фильтрации по тестовому сигналу в виде суммы гармоник.

Ключевые слова: цифровой фильтр, аналоговый прототип, цифровой сигнальный процессор, синусоидальный сигнал, фильтр нижних частот.

Система, выполняющая математические операции с дискретизированными сигналами для уменьшения или выделения определенных частот сигнала, называется цифровым фильтром. Цифровая фильтрация — это программная реализация алгоритма. Как правило, аппаратная часть состоит из аналого-цифрового преобразователя (АЦП) для оцифровывания входного сигнала, микропроцессора и периферийных компонентов, таких как память для хранения данных и коэффициентов фильтра и т.д. Выделяются два обширных класса: фильтры с бесконечной импульсной характеристикой (БИХ-фильтры) и фильтры с конечной импульсной характеристикой (КИХ-фильтры). Каждый из типов (в стандартной форме) можно представить через коэффициенты его импульсной характеристики $h(s)$ (реакция фильтра на единичный импульс при нулевых начальных условиях). Выбор между КИХ и БИХ фильтрами зависит от относительных преимуществ обоих типов. КИХ-фильтры всегда устойчивы и могут иметь строго линейную фазовую характеристику. Из этого следует, что фильтр не вводит фазового искажения в сигнал, что важно во многих сферах, например, обработке данных, биомедицине, цифровой аудиообработке или обработке изображений. Импульсная характеристика БИХ-фильтров продолжается бесконечно, что при неправильном проектировании может привести к неустойчивости фильтра. Основным преимуществом БИХ-фильтров является эффективность в реализации. Порядок фильтра БИХ может быть значительно ниже, а это уменьшает количество необходимых ресурсов и соответственно требования к аппаратной части по сравнению с КИХ фильтрами.

Как правило, проектирование цифрового фильтра начинается с исследования аналогового фильтра прототипа (фильтры Чебышева, эллиптические фильтры и т.д.). Хорошо изучена теория этих фильтров. Основной их характеристикой является передаточная функция $K(s)$, которая может быть полностью описана набором нулей и полюсов. Эти нули и полюса отображаются на s -плоскости. По АЧХ фильтра может быть определена т.н. частота среза, разделяющая полосу пропускания от полосы задержания.

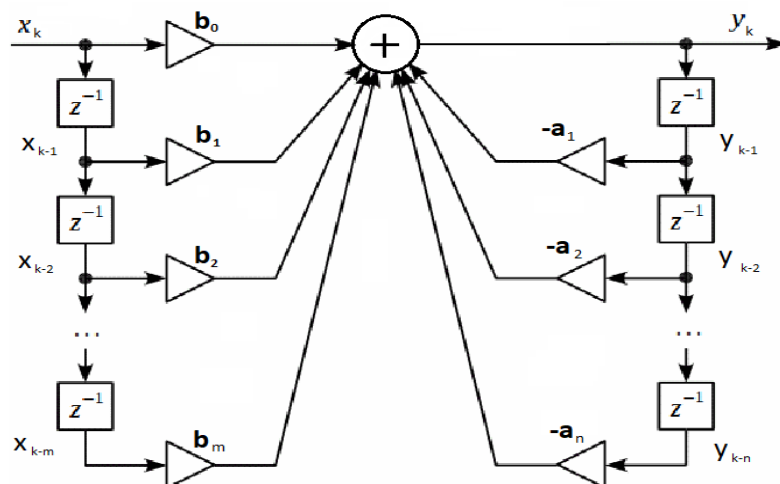


Рис. 1. Структурная схема цифрового фильтра

При проектировании цифрового фильтра производится т.н. билинейное z -преобразование (переход из s -плоскости в z -плоскость) в результате которого могут быть получены коэффициенты передаточной характеристики в z -плоскости и получены коэффициенты a_k и b_k в модели реализации цифрового фильтра.

Для обработки оцифрованных сигналов применяются специализированные цифровые сигнальные процессоры (от англ. Digital Signal Processor, DSP) [1]. Цифровая фильтрация - лишь одно из направлений, для которого применяются цифровые сигнальные процессоры, так же они выполняют такие задачи, как преобразование Фурье, поиск сигналов и т.п. Математически эти задачи сводятся к поэлементному перемножению элементов многокомпонентных векторов действительных чисел, последующему суммированию произведений (например, в цифровой фильтрации выходной сигнал фильтра с конечной импульсной характеристикой равен сумме произведений коэффициентов фильтра на вектор выборок сигнала, аналогичные вычисления производятся при поиске максимумов корреляционных и автокорреляционных функций выборок сигналов). Именно поэтому цифровые сигнальные процессоры оптимизированы по быстродействию для выполнения именно таких операций. По типу арифметики ЦСП делятся на процессоры, которые обрабатывают данные с плавающей точкой и с фиксированной точкой. Устройства с плавающей точкой позволяют производить вычисления с большей точностью (с меньшей погрешностью коэффициентов фильтра), но, в свою очередь, их цена на порядок выше из-за более сложного устройства. Так же уменьшается быстродействие. Как правило, ЦСП с фиксированной точкой обрабатывают 16-ти разрядные данные, а процессоры с плавающей точкой данные с разрядностью 32 бита.

С помощью DSP System Toolbox (MatLab) можно проектировать и анализировать разнообразные фильтры, в том числе фильтры с КИХ и БИХ. Можно передавать сигналы от переменных, файлов и сетевых устройств для разработки и проверки системы. Временные рамки, анализатор спектра позволяют динамически визуализировать и измерять потоковые сигналы. Для настольного использования и развертывания на встроенных процессорах, таких как ARM Cortex system toolbox поддерживает генерацию кода на языке C / C++. Алгоритмы доступны в виде функций MATLAB, системных объектов и блоков Simulink. Блоки Simulink для обработки сигналов поддерживают типы данных двойной точности и одинарной точности с плавающей точкой, с фиксированной точкой, а также целочисленные типы данных.

Для проектирования цифрового фильтра используется блок Simulink под названием Filter Designer. Этот блок представляет большие возможности для проектирования и настройки фильтра. В этом окне есть возможность выбрать один из различных БИХ или

КИХ фильтров, задать предполагаемую частоту дискретизации входного сигнала, который планируется фильтровать, задать значения полосы пропускания и определиться с тем, фильтр каких частот будет рассчитан. Так же можно задать конкретный или рассчитать минимально возможный порядок фильтра. Например, фильтр Хэмминга нижних частот с конечной импульсной характеристикой. Для этого в блоке «Design Method» необходимо поставить чек-бокс возле поля «FIR» (finite impulse response) и в выпадающем списке выбираем «Window». Далее в блоке Options выбираем интересующий нас фильтр «Hamming». В дальнейшем используется синусоидальный сигнал с частотой дискретизации 48000. Имея это ввиду, задается частота дискретизации в соответствующем поле Filter Designer. Полоса пропускания задается равной 2000 Гц. Для сравнения задается порядок фильтра, равный пятидесяти, соответственно это необходимо указать в поле «Specify order». Для того, чтобы коэффициенты фильтра рассчитались после ввода всех исходных данных, нужно нажать на кнопку «Design Filter». На современных компьютерах расчет фильтра такого порядка занимает в среднем не более тридцати секунд, а это время может уменьшиться при использовании более современной и мощной аппаратной части.

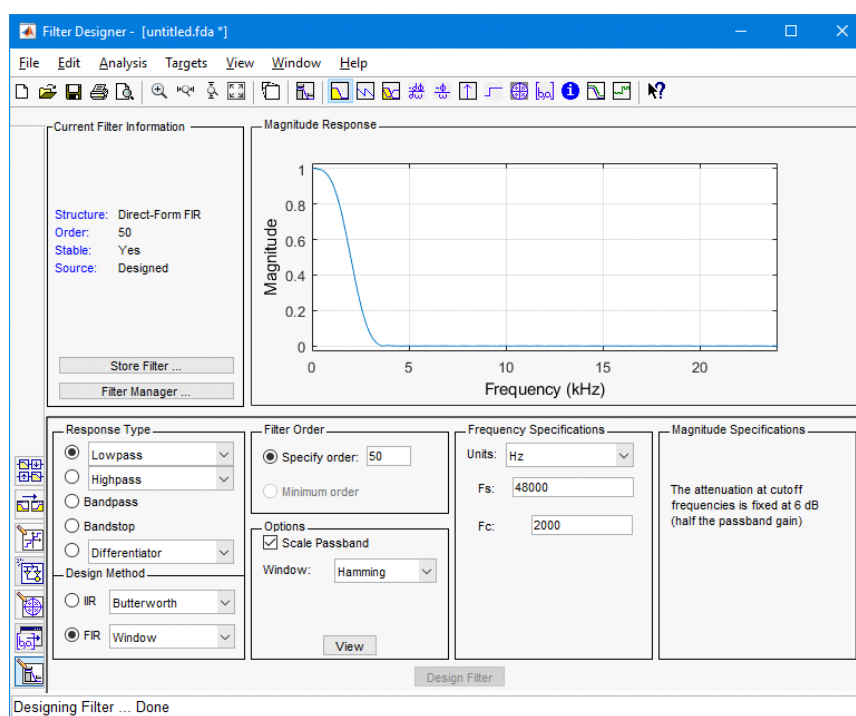


Рис. 2. Окно Filter Designer с АЧХ проектируемого фильтра Хэмминга 50го порядка

Представлена амплитудно-частотная характеристика (АЧХ). По оси ординат отношение амплитуд на входе и выходе фильтра может откладываться в абсолютных значениях либо в децибелах, сделать это можно, нажав ПКМ по надписи «Magnitude (dB)» расположенной левее графика. После изменения график становится более понятным для восприятия.

После успешного расчета коэффициентов фильтра можно использовать их в реализации фильтра. Поскольку в дальнейшем будет использоваться DSP, то лучшим вариантом будет экспорт коэффициентов в файл [2]. Это можно сделать в контекстном меню «Targets», в котором необходимо выбрать пункт «Generate C header ...». Следом нажимать «Generate». Нужно указать путь для сохранения файла в формате «.h». Именно в нем и будут храниться коэффициенты фильтра пятидесятого порядка. Далее необходимо перенести эти коэффициенты любым способом в CCS.

Выгруженные из MATLAB коэффициенты, можно открыть с помощью редактора Notepad++, после чего можно перенести их в CCS (Code Composer Studio).


```

const int BL = 51;
const real64_T B[51] = {
0.0002625788399601,-2.705769969277e-19,-0.0003369705144782,-0.0008091918280468,
-0.001461434207739,-0.002304311277159,-0.003298885835117,-0.004346700411715,
-0.005287441678118,-0.005905562164873,-0.005946056128625,-0.005138358180181,
-0.003226162047124,1.840630295005e-18,0.004671183606732,0.01081339368463,
0.01832948676811,0.02699243250541,0.03645157958487,0.0462521452293,
0.05586663501279,0.06473548345036,0.07231306023298,0.07811445697661,
0.08175825130394,0.08300077415495,0.08175825130394,0.07811445697661,
0.07231306023298,0.06473548345036,0.05586663501279,0.0462521452293,
0.03645157958487,0.02699243250541,0.01832948676811,0.01081339368463,
0.004671183606732,1.840630295005e-18,-0.003226162047124,-0.005138358180181,
-0.005946056128625,-0.005905562164873,-0.005287441678118,-0.004346700411715,
-0.003298885835117,-0.002304311277159,-0.001461434207739,-0.0008091918280468,
-0.0003369705144782,-2.705769969277e-19,0.0002625788399601
};

```

Рис. 3. Коэффициенты проектируемого фильтра

Для генерации сигнала применяется программ CoolEdit. Этот инструментарий используется при создании и обработки аудиофайлов. Позднее программа получила новое название Adobe Audition. Внимание пользователей CoolEdit заслужил благодаря возможности работы с большим количеством аудиодорожек и с файлами большой длины. Несмотря на то, что последний релиз программы был довольно давно, она имеет похожий на современный интерфейс. Приложение подходит для нарезки, сведения и склейке аудиофайлов. Для дальнейшей работы потребуется сгенерировать синусоидальные сигналы различной частоты и сложить их. Поскольку частота среза была выбрана равной 2000 Гц, то генерируемый сигнал, выбираем состоящим из гармоник 1000 Гц и 3000 Гц (Рис. 4). Для генерации необходимо переключиться в режим Edit View, нажав на соответствующую кнопку в верхнем левом углу. Далее следует нажать на Generate и в выпадающем меню нажать кнопку Tones. После открывается окно, в котором необходимо указать частоту дискретизации, разрядность и выбрать mono или stereo сигнал будет генерироваться.

Аудиокодек aic3204 содержит 16-ти битный АЦП, конфигурируется и настраивается внешним процессором по интерфейсу I2C. Запускается функция, которая передает два значения: 7-ми битный адрес устройства и 8-ми битные данные регистра. Затем эти данные передаются в функцию для подсчета количества байт для записи. Это происходит, когда I2C сконфигурирован как ведущий передатчик. Далее в регистр I2C_SAR передается 7-ми разрядный адрес подчиненного устройства. За тем I2C переводится в режим мастера записи. После короткой паузы начинается процесс записи данных. CPU записывают данные до 8 бит, если данные меньше 8 бит, то они выравниваются. Во время передачи данных используется скоростной интерфейс I2S. После завершения передачи данных в отведенный регистр посылается сигнал STOP.

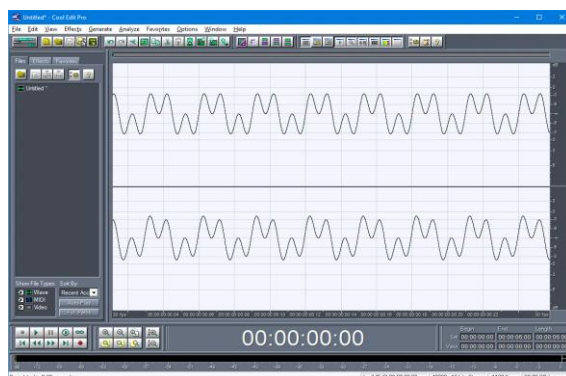


Рис. 4. Сигнал 1000 Гц + 3000 Гц

Данный аудиокодек размещён на отладочной плате USB Stick 5535 и с помощью него производится оцифровка сигнала и передача его в процессор.

Для дальнейшей работы с аудиокодеком aic3204 нам необходим стереосигнал с частотой дискретизации 48000. Следом открывается окно, в котором задаются частотные характеристики самого сигнала. Таким образом, сигнал реализован из двух гармоник с разной амплитудой, путем выставления коэффициентов умножения.

На данный момент разработано программное обеспечение для фильтрации сигналов на DSP TI TMS320C5535 [3]. Этот 32-х битный микропроцессор предназначен для работы с цифровыми сигналами. Тактовая частота процессора до 120 МГц, 320 Кбайт RAM и 128 Кбайт ROM. Архитектура C55xmkm обеспечивает высокую производительность и низкое энергопотребление за счет параллелизма вычислений и особого внимания к энергопотреблению. Процессор поддерживает определенную внутреннюю структуру шин, которая состоит из одной программной шины, одной 32 битной шины чтения данных, двух 16 битных шин чтения данных, двух 16 битных шин записи данных и дополнительных шин для работы с периферией и DMA (Direct Memory Access или иначе ПДП - Прямой Доступ к Памяти). Данная конфигурация шин позволяет производить до четырех операций чтения 16 битных данных и до двух операций записи 16 битных данных за один цикл. Процессор C55x имеет два устройства для выполнения операции умножения с накоплением MAC (multiply-accumulate), каждое из которых может выполнить умножение чисел семнадцать на семнадцать бит и суммирование с тридцатью двумя битами за один цикл. Так же имеется 40-битное арифметико-логическое устройство (АЛУ) и дополнительное к нему 16-ти битное АЛУ. На «борту» процессора имеется 10 битный АЦП последовательного приближения. Последовательные media интерфейсы поддерживают два типа периферийных устройств MultiMedia Card/Secure Digital (MMC/SD), четыре Inter-IC Sound (I2S шина) модуля, один Serial-Port Interface (SPI), один I2C multi-master and slave интерфейс и универсальный асинхронный приемник/передатчик (UART). Процессор поддерживается средой разработки Code Composer Studio, может работать под управлением операционной системы OS Real Time DSP/BIOS, поддерживает высокоскоростной обмен данными RTDX (Real Time Data eXchange) с компьютером по JTAG интерфейсу XDS100.

Code Composer Studio - это мощная интегрированная среда разработки для создания кода для DSP и ARM процессоров семейства TMS320 и других процессоров выпущенный компанией Texas Instruments. Создание и отладка проекта в Code Composer Studio происходит в стандартной оболочке Eclipse. Она имеет 2 окна: C++/C и Debug. В коде фильтра присутствует функция aic3204_codec_read (&left_input, &right_input). Поскольку каждый семпл сигнала обрабатывается по отдельности, мы сначала считываем подаваемый цифровой сигнал этой функцией. Происходит это путем передачи в нее переменных левого и правого канала входного сигнала, которым присвоены значения регистров процессора I2SRXLT1 и I2SRXRT1 (Receive Left/Right Data 1 Register). Далее полученные данные для левого и правого каналов преобразуются в один с при помощи функции stereo_to_mono(left_input, right_input). В этой функции находится среднее арифметическое между значениями левого и правого каналов.

После переноса коэффициентов фильтра в исходный код необходимо загрузить его на отладочную плату и подать на ее вход сгенерированный в CoolEdit сигнал. Увидеть результат работы фильтра можно с помощью стандартных функций CCS. Для построения графика необходимо создать необходимое количество массивов и передать в него результаты данные (входной оцифрованный сигнал и сигнал после фильтрации). Далее во вкладке Tools выбрать пункт Graph и в выпадающем списке Single Time. В открывшемся окне в поле Start Address указать адрес созданного массива, так же указать размер массива и в поле Dsp Data Type выбрать 16 bit signed integer.

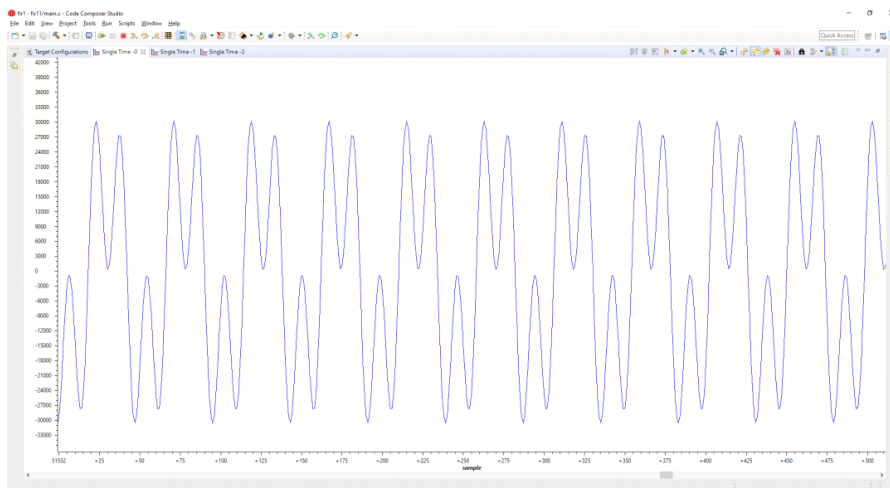


Рис. 5. Сигнал 1000 Гц + 3000 Гц в CCS

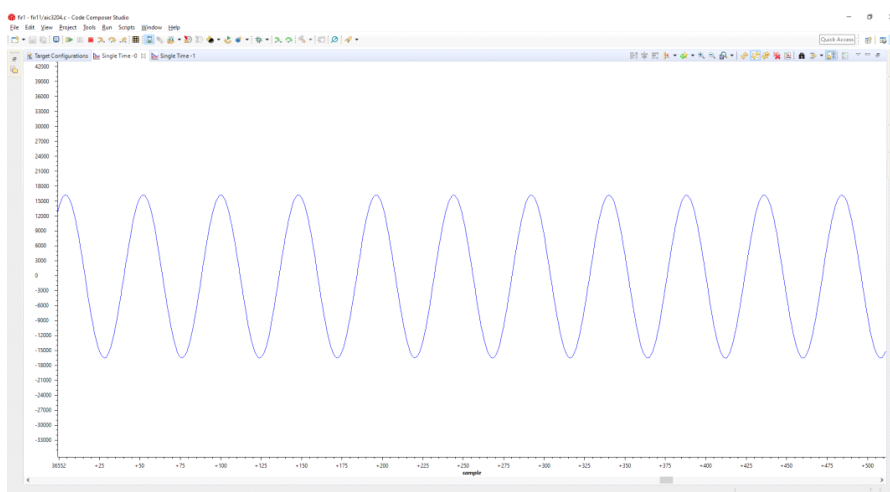


Рис. 6. Сигнал после фильтрации

Результатом данной работы явилось исследование алгоритма цифровой фильтрации, изучение и использование целевых ресурсов в среде MatLab, программная реализация на цифровом сигнальном процессоре с использованием среды разработки Code Composer Studio.

Библиографический список

1. Сергиенко А.Б. Цифровая обработка сигналов // Питер. -2003. -С.768
2. Добровольский Р.В. Цифровая фильтрация аудио сигналов в режиме реального времени на базе процессора TMS320C5515 / Р.В. Добровольский, Т.А. Ким, А.А. Сотников // Технологии инженерных и информационных систем. -2019. -№1. -С.11-20.
3. TMS320C5515 eZdsp Usb Stick. Texas Instruments, 2010. - URL: http://support.spectrumdigital.com/boards/usbstk5515/revA/files/usbstk5515_TechRef_RevA.pdf

D.E. Kogan, O.N. Korelin

DIGITAL FILTERING USING THE TI 5535 DSP

Nizhniy Novgorod state technical university n. a. R. E. Alekseev

Digital filtering is one of the most powerful tools in digital signal processing. In addition to the obvious advantages, digital filters can meet technical requirements that would be extremely difficult to achieve in an analog version. The article discusses the methods of designing and applying a digital filter using a digital signal processor. The coefficients are calculated and imported into the program code, a sinusoidal signal of various harmonics is created, followed by filtering.

Keywords: digital filter, analog prototype, digital signal processor, sinusoidal signal, low-pass filter.

АЛГОРИТМ ПОИСКА ЗАИМСТВОВАНИЙ В ТЕКСТАХ НА РУССКОМ ЯЗЫКЕ С УЧЕТОМ ИХ СЕМАНТИЧЕСКИХ И СИНТАКСИЧЕСКИХ ОСОБЕННОСТЕЙ

Нижегородский государственный технический университет им. Р.Е. Алексеева

Проблема плагиата стоит наиболее остро в научном сообществе. Алгоритмы, используемые в системах обнаружения плагиата, являются несовершенными, и их недостатки пытаются исправить многие исследователи. При этом в основном решается задача обнаружения плагиата текста, а не смысла. В данной работе предлагается нетипичный подход поиска заимствований в тексте, основанный на моделировании семантического и синтаксического уровней языковой системы внутри коллекции текстов академической тематики в области IT. Предполагается, что данный подход улучшит качество нахождения заимствований и будет способствовать изменению подхода при написании научных работ.

Ключевые слова: плагиат, моделирование семантического уровня, моделирование синтаксического уровня, перефразирование текста на русском языке, языковые модели, ANN (ANNOY).

Введение

В современном мире в связи с быстрым развитием сети Интернет плагиат проникает во многие сферы человеческой жизни: научная сфера, образование, промышленность, множество творческих областей и многое другое. В 2009 году была опубликована статья [1], где были приведены результаты социологического опроса в НИУ ВШЭ по проблеме плагиата. Около 50 % студентов российских вузов скачивают готовые работы (эссе, рефератов, курсовых) и выдают их в качестве своих. Наиболее остро проблема плагиата стоит в научном сообществе и сфере промышленности. В 2013 году компания "Анти-плагиат" по заказу российской государственной библиотеки проводила исследование, в результате которого оказалось, что в 10% диссертаций, защищенных в России с 2000 года, количество плагиата превышало 70%. "Наша задача была - оценить размер бедствия. Размеры бедствия огромные" – заявил директор российской государственной библиотеки Александр Вислый [2].

Обзор функционирования системы обнаружения плагиата

Для того, чтобы работа соответствовала заявленным нормам по критерию плагиата, автору работы необходимо проверить текст на наличие плагиата перед публикацией или загрузкой в систему. В настоящий момент существуют разнообразные системы обнаружения плагиата. Чтобы снизить уязвимость такой системы, коммерческие сервисы определения плагиата (СОП) не раскрывают алгоритмов, которые используются при проверке. В статье [3] приводится типовая структура СОП, которая была получена путем анализа открытых СОП и рассмотрения декларируемых принципов работы СОП. Она включает в себя следующие блоки: загрузка текста в систему, очистка текста от исходного форматирования, передача очищенного текста в качестве запроса в базу данных СОП, получение набора документов – вероятных источников плагиата. Кроме того, такая СОП, по соглашению с пользователем, создает копию документа, которую может использовать для дальнейших проверок. Данный аспект не корректен по отношению к текстам, где требуется соблюсти нормы конфиденциальности загружаемого на проверку документа.

К преимуществам большинства существующих систем можно отнести то, что крайне сложно придумать эффективный способ обмана такой системы. Но подобные интернет-версии зачастую предоставляют услуги на платной основе, количество единиц

текста на проверку может быть ограничено, часто поиск осуществляется продолжительное время [4]. Кроме того, у подобных алгоритмов отсутствует возможность настройки алгоритма на конкретную специфическую область.

Актуальность проблемы определения плагиата, в том числе смыслового, инициирует поиск новых решений в этой области. В работе [3] предлагается новая архитектура системы СОП, которая должна улучшить производительность СОП, а также позволить искать совпадения текста в сети интернет, при этом загружаемый документ для сторонней организации будет представлен в «нечитабельном» виде. В работе [4] сделана попытка разработки системы проверки авторства текста с учетом ее специфической области. Для этой цели автор статьи сравнивает результаты применения ряда алгоритмов для решения поставленной задачи, а именно исследует метод шинглов, метод сравнения подмножества слов без пересечения, метод сравнения 30% наиболее встречающихся слов, а также комбинацию метода шинглов и метода сравнения наиболее встречающихся слов. Под шинглом в рассматриваемой работе понимается последовательность слов в документе, причем выборка данных последовательностей осуществляется «внахлест», а не «в стык». В работе показано, что лучшее качество проверки текста на антиплагиат было получено при использовании метода шингла (такое же качество наблюдалось при использовании системы antiplagiat.ru), однако этот метод оказался самым длительным по реализации. При комбинации двух методов значение показателя качества поиска плагиата такое же, как и для метода шингла, но время выполнения сократилось на 60%, в связи с чем автор работы [4] принял решение использовать данную комбинацию методов для разработки системы.

Анализ работ в исследуемой области показал, что при улучшении работы СОП во внимание не берется семантический аспект текста. Значение термина «плагиат» в существующих решениях, как правило, означает прямое копирование текста, а не плагиат идей, когда автор документа заменяет слова на синонимы и изменяет, таким образом, синтаксическую структуру предложения.

Постановка задачи исследования

Целью данной работы является исследование и реализация нового подхода к решению задачи поиска заимствований в тексте, когда плагиат рассматривается со смысловой точки зрения, при этом будут использоваться не последовательности слов, а семантическая составляющая частей текста и синтаксический строй предложения.

Разрабатываемый алгоритм должен реализовать поиск схожих текстов в заданной исследовательской базе. К рассмотрению будем принимать научные тексты, диссертации, курсовые, где уникальность работы и новизна являются ключевым фактором. Ставится задача усовершенствовать алгоритм «Антиплагиат» таким образом, чтобы при изменении синтаксического дерева предложения алгоритм был способен установить, из какого источника произошло заимствование, и мог рассчитать процент уникальности текста. Решение должно вначале производить анализ текста сравнением по предложениям текстов коллекции, а затем перейти к анализу смыслового содержания текстов.

На данный момент для русского языка не существует корпусов, позволяющих сравнивать смысловую схожесть текстов. Для сравнения предложений по смыслу ситуация немного лучше: в свободном доступе удалось найти небольшие коллекции перефразированных предложений [5]. В связи с этим возникает задача создания коллекции перефразированных текстов для тестирования работы разработанной системы. Для этих целей была разработана отдельная программа, которая с помощью изменения синтаксической структуры предложения, т.е. перестановкой различных членов предложения внутри предложения или разделением сложного предложения на простые и т.д., а также с помощью частичной или полной замены значимых слов на синонимы, позволяет перефразировать исходное предложение, а затем и весь текст целиком.

Описание составляющих разрабатываемой системы

В качестве *гипотезы*, выносимой на проверку в рамках данного исследования, берется следующее положение: при эксплицитном задании синтаксического дерева предложения и установления отношения слово-синтаксическая роль алгоритм проверки текста на уникальность будет показывать лучшую точность. То есть, если предположить, что изначально алгоритм сравнения двух предложений между собой явно не учитывает изменения синтаксического дерева, то он и не сможет установить в полной мере схожесть этих предложений. Общее представление такого алгоритма представлено на рис. 1



Рис. 1. Общее представление алгоритма поиска схожести двух предложений

В данном решении важную роль играет способ векторизации предложения на входе. Так, существуют классические подходы: TF-IDF, Doc2vec, которые позволяют представить предложение в виде вектора, но при этом не учитывается связь слов в предложении.

Новыми технологиями в векторном представлении предложений считаются модели *InferSent* и *sentence-transformer* (BERT), которые являются предобученными лингвистическими моделями, которые особенны тем, что при построении вектора учитывают взаимосвязи слов в предложении. Однако работа предобученных моделей является некоторым «черным ящиком» и явно сказать, как именно происходит моделирование синтаксического уровня, не предоставляется возможным. Отсюда и возникла идея явного задания синтаксической структуры. Предполагается, что данный подход позволит улучшить качество работы алгоритма. Вектор синтаксического дерева предлагается соединить с вектором, полученным с помощью предобученной модели.

Для оценки схожести двух предложений используют разные способы: определение косинусного расстояния, метод *k* ближайших соседей (KNN) – алгоритм HNSW (библиотека *hnsplib*), метод приближенного поиска ближайших соседей (ANN) – использование библиотеки ANNOY [6].

В рамках данной работы были проведены эксперименты с использованием метода определения косинусного расстояния и метода приближенного поиска ближайшего соседа библиотеки ANNOY на 10000 примерах. В отличие от точного поиска ближайшего соседа,

в методе приближенного поиска ближайших соседей ANN внутри алгоритма данные разделяются на более мелкие части схожих векторных представлений. Поиск по таким частям осуществляется быстрее и эффективнее.

Результаты произведенных экспериментов для сравнения предложений

Для выбора оптимального способа векторизации и исследования алгоритма сравнения схожести двух предложений были произведены следующие эксперименты:

- были использованы две модели BERT: distilbert-base-nli-stsb-mean-tokens и distiluse-base-multilingual-cased-v2. Эти модели BERT выбраны из-за их возможности работать с русскоязычными данными;
- проведено исследование двух методов сравнения двух предложений – метода определения косинусного расстояния и метода приближенного поиска ближайших соседей использование ANN библиотеки ANNOY;
- в качестве данных, на которых производилось сравнение, были взяты 10000 примеров перефразированных предложений [5].

В табл. 1 представлены некоторые примеры данных, с которыми производились эксперименты. Для каждого предложения из колонки sentence_1 находилось 5 близких по смыслу предложений и затем проверялось, имеется ли в данном наборе индекс истинно схожего предложения (предложение с тем же индексом, что и рассматриваемое, но из колонки sentence_2). Если в найденном наборе предложений присутствовал индекс истинного предложения, то показатель числа верно найденных предложений увеличивался на единицу. Показатель качества работы алгоритма рассчитывался как отношение числа верно найденных предложений к количеству всех предложений.

Таблица 1

Структура использованных данных в рамках эксперимента

	Предложение_1 (sentence_1)	Предложение_2 (sentence_2)
0	звоните ответим на все вопросы	звоните на все вопросы ответим
1	не поет и не танцует	не танцует и не поет
2	автомобиль приобретался у официального дилера	приобретался автомобиль у официального дилера
3	все вопросы на почту пожалуйста	все вопросы пожалуйста на почту
4	пишите звоните прямо сейчас	звоните пишите прямо сейчас

В результате были получены относительные частоты правильно найденных схожих предложений применением исследуемых методов BERT, представленные в табл. 2.

Таблица 2

Результаты экспериментов для сравнения предложений

	Определения косинусного расстояния Cos_sim	Приблизительный поиск ближайшего соседа (ANN)
distilbert-base-nli-stsb-mean-tokens	0.94	0.87
distiluse-base-multilingual-cased-v2	0.992	0.994

Использование алгоритма ANN дает большое преимущество по скорости работы. При этом оба алгоритма дают практически одинаковые значения показателей точности для модели «distiluse-base-multilingual-cased-v2». Таким образом, для дальнейшей разработки системы будет использован алгоритм ANN.

В данном исследовании при построении системы поиска заимствований в русскоязычных текстах с учетом их семантических и синтаксических особенностей используется модель BERT distiluse-base-multilingual-cased-v2, которая предобучена на общих текстах. В дальнейшем, при переходе к исследованиям текстов в IT-области,

необходимо будет произвести обучение модели BERT на текстах академической тематики с уклоном в область применения информационных технологий.

Также важно отметить, что поиск соответствий будет осуществляться внутри некоторой коллекции текстов заданной тематики, а не текстов из сети Интернет. Это обусловлено тем, что суть алгоритма состоит в моделировании семантического и синтаксического уровней языка, а моделирование данных уровней для текстов из сети Интернет может в значительной степени замедлить скорость работы алгоритма.

Описание общего принципа работы разрабатываемой системы

Основная идея разрабатываемой системы представлена на рис. 2.

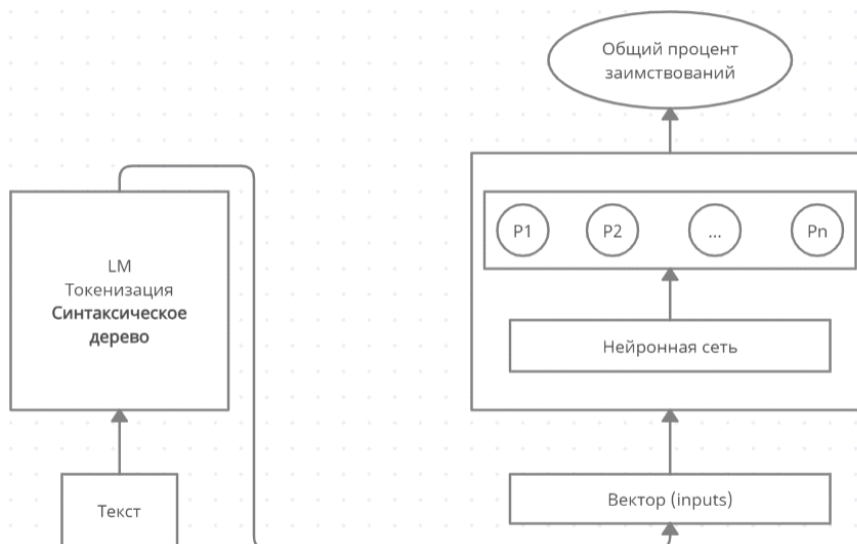


Рис. 2. Общий алгоритм работы системы

Суть процесса заключается в следующем: сначала текст, который требуется проверить на плагиат, поступает на предобработку, токенизируется, строится вектор входного текста. Вектор входного текста представляет собой объединение векторов составляющих текст предложений. Вектор предложения, как было отмечено ранее, получается в результате объединения вектора «смысла» - вектора, полученного в результате использования модели BERT, и вектора представления синтаксической структуры предложения. Затем полученный вектор подается на вход нейронной сети. В дальнейшем нейронная сеть будет обучена на текстах научной тематики – сопоставление перефразированных текстов и текстов без перефразирования, т.е. оригинальных источников. На выходе нейронной сети будет получен процент отнесенности текста к каждому тексту из базы научных текстов. На основании полученных результатов будет производиться определение общего процента заимствований.

Заключение

В статье рассмотрен нетипичный подход для решения задачи поиска заимствований в тексте, заключающийся в моделировании семантического и синтаксического уровня языка для входного текста. Поиск заимствований предлагается осуществлять внутри коллекции документов заданной специфической области. В качестве алгоритма сравнения двух предложений был выбран алгоритм ANN (библиотека ANNOY). Для получения векторного представления предложения и, тем самым, задания семантики предложения, была выбрана модель BERT.

В дальнейшем планируется обучить BERT модель на текстах академической тематики в области применения IT-технологий, что является, безусловно, актуальным, а также применить подходы, рассматриваемые в рамках предложений, на текст целиком.

Библиографический список

1. Ивойлова И. Украденные мысли: половина студенческих рефератов и курсовых скачивается из Интернета // Российская газета. 2009. No 4830. 20 янв. URL: <http://www.rg.ru/2009/01/20/referaty.html>
2. 10% диссертаций по истории, защищенных с 2000 г., оказались плагиатом // Ведомости. 2013 URL: <https://www.vedomosti.ru/management/articles/2013/09/04/v-dissertatsiyah-po-istorii-zaschischennyh>
3. Дягилев В.В., Цхай А.А., Бутаков С.В. Архитектура сервиса определения плагиата, исключая возможность нарушения авторских прав (рус.) // Вестник НГУ. Серия: Информационные технологии. — 2011. — С. 23—29.
4. Зиберт А.О., Хрусталеv В.И. Разработка системы определения наличия заимствований в работах студентов высших учебных заведений. Алгоритмы поиска нечетких дубликатов // Universum: Технические науки: электрон. научн. журн. 2014. No 3 (4).
5. Коллекции перепарафразированных предложений // URL: https://github.com/Koziev/NLP_Datasets/tree/master/ParaphraseDetection
6. Документация ANNOY // URL: <https://github.com/spotify/annoy>

E.S. Sokolova, A.D. Tryaskova

THE ALGORITHM FOR FINDING LOANWORDS IN TEXTS IN RUSSIAN WITH THEIR SEMANTIC AND SYNTACTIC FEATURES

Nizhny Novgorod state technical university. R.E. Alekseeva

Due to the rapid development of the Internet, the problem of plagiarism is most acute in the scientific community. The algorithms used in plagiarism detection systems are not perfect, and many researchers are trying to correct their shortcomings. However, nowadays the task of text plagiarizing, is mostly solved. The meaning plagiarizing is quite unsolved. In the current research, there is proposed an atypical approach to the search for loanwords in the text, based on modeling the semantic and syntactic levels of the language system. The correlation is searched within the academic text in the IT field. It is assumed that this approach will improve the quality of finding loanwords and will contribute to changing the approach for writing scientific papers.

Keywords: plagiarism, semantic level modeling, syntactic level modeling, paraphrasing of text in Russian, language models, ANN (ANNOY)

СЕКЦИЯ 5.2 ТЕХНИЧЕСКАЯ КИБЕРНЕТИКА (ИНТЕЛЛЕКТУАЛЬНЫЕ СИСТЕМЫ УПРАВЛЕНИЯ)

УДК 004.021

В.А. Чеканин

РЕШЕНИЕ ЗАДАЧИ РАЦИОНАЛЬНОЙ КОМПОНОВКИ ДЕТАЛЕЙ СЛОЖНОЙ ГЕОМЕТРИЧЕСКОЙ ФОРМЫ

Московский государственный технологический университет «СТАНКИН»

Цель работы: Разработка алгоритма, обеспечивающего быстрое решение задачи плотной компоновки деталей сложной геометрической формы.

Методика: Первоначально детали сложной геометрической формы подвергаются вокселизации, обеспечивающей их перевод в дискретное представление, далее выполняется декомпозиция полученных ортогональных многогранников для сокращения числа составляющих их ортогональных объектов, после чего применяется разработанный алгоритм размещения ортогональных многогранников произвольной размерности.

Результаты: Высокая эффективность применения разработанных алгоритмов формирования и размещения ортогональных многогранников продемонстрирована на примерах решения задач компоновки деталей на платформе 3D-принтера.

Выводы: Реализованный подход к решению задач рационального размещения деталей сложной геометрической формы обеспечивает быстрое формирование плотных компоновок. Разработанный алгоритм размещения ортогональных многогранников осуществляет выбор наиболее подходящих свободных областей контейнера для получения наиболее компактной компоновки деталей.

Ключевые слова: компоновка деталей, задача упаковки, ортогональный многогранник, аддитивное производство.

Введение

Решение проблемы оптимизации компоновки деталей имеет большое значение в аддитивном производстве, поскольку оно позволяет существенно сократить расход материала, уменьшить время как на подготовку схемы размещения деталей внутри платформы (контейнера) 3D-принтера, так и на их последующее непосредственное изготовление. Использование более плотной компоновки дополнительно приводит к сокращению энергетических затрат в процессе изготовления объектов, а также снижает износ используемого оборудования. Основные трудности, возникающие при решении этой задачи, связаны как со сложной формой размещаемых объектов, так и с её принадлежностью к классу NP-трудных оптимизационных задач [1–3].

Для задач фигурного раскроя и упаковки геометрически сложных объектов, заданных в полигональном виде, наиболее распространён способ определения взаимного положения объектов с использованием годографа функции плотного размещения, предложенного Ю.Г. Стояном [4, 5]. Годограф функции плотного размещения объекта определяет границу области, которая не может быть использована для размещения этого объекта. В процессе поиска области для размещения объекта в контейнере выполняется моделирование его движения относительно других размещённых в контейнере объектов [5] и определение точек пересечения годографов функций, составленных для размещаемого объекта и всех находящихся в контейнере объектов [6]. Важно отметить, что для размещения нового типа объекта требуется либо выполнить его декомпозицию на набор базовых геометрически простых фигур, для которых существует соответствующая математическая модель (в виде Φ -функции), либо разработать новую математическую модель для конкретного объекта с кусочно-гладкими границами [7], что в конечном итоге

ограничивает области практического применения этого способа для деталей сложной формы. Другим недостатком этого способа является необходимость применения методов нелинейного программирования, требующих решения системы большого числа неравенств [5], которые на практике при увеличении числа размещаемых объектов оказываются неприменимыми из-за значительных затрат временных ресурсов.

Указанные недостатки имеют место при использовании полигонального представления объектов, однако они отсутствуют при использовании воксельной модели, обеспечивающей геометрическое описание объектов и контейнеров произвольной формы и произвольной размерности в виде конечного набора дискретных элементов (вокселей) [8, 9], который в ряде работ называется также рецепторной матрицей [10].

Методика

Для размещения деталей сложной геометрической формы в настоящей работе предлагается подход, заключающийся в преобразовании объектов в воксельное представление и последующем применении разработанного алгоритма размещения ортогональных многогранников (рис. 1).

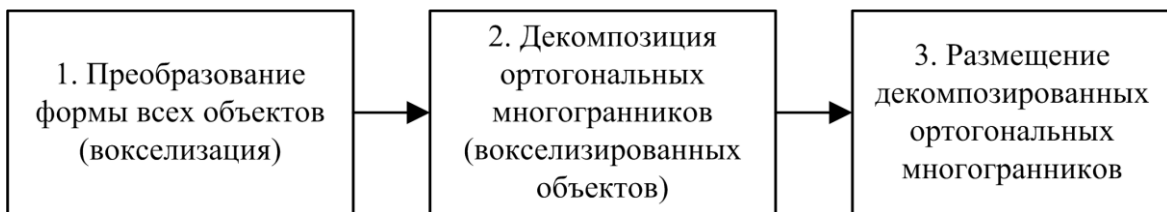


Рис. 1. Реализованный подход к решению задач компоновки

Под ортогональным многогранником (ОМ) O_i размерности D понимается набор из m_i ортогональных объектов $o_{i,k}$, $k \in \{1, \dots, m_i\}$ в виде D -мерных параллелепипедов с фиксированным положением друг относительно друга [11, 12].

В результате вокселизации каждая размещаемая деталь будет представлена в виде ортогонального многогранника, состоящего из большого числа одинаковых D -мерных кубов заданного размера. Для повышения скорости формирования компоновки выполняется декомпозиция всех ортогональных многогранников, в результате которой сокращается число ортогональных объектов, описывающих ортогональные многогранники за счёт объединения вокселей в крупные ортогональные объекты [13].

В основе предложенного алгоритма компоновки деталей сложной формы лежит использование модели потенциальных контейнеров. Она описывает свободное пространство заполняемого объектами контейнера в виде набора всех возможных ортогональных областей с максимальными габаритными размерами (т.н. потенциальных контейнеров, ПК), которые могут быть использованы для последующего размещения в них объектов [14, 15].

На рис. 2 представлена блок-схема разработанного алгоритма размещения в контейнере ортогонального многогранника O_i произвольной размерности. Здесь используются следующие обозначения: $U_{i,k}$ – ортогональный многогранник, определяющий область допустимого размещения ортогонального многогранника O_i в текущем ПК при размещении объектов $o_{i,1}$ и $o_{i,k}$, U_i – ортогональный многогранник, определяющий область допустимого размещения ортогонального многогранника O_i в текущем ПК. В процессе определения области допустимого размещения применяется операция пересечения ортогональных многогранников, результатом выполнения которой является новый ортогональный многогранник, точки которого занимают пространство, принадлежащее одновременно двум ортогональным многогранникам, участвующим в этой операции.

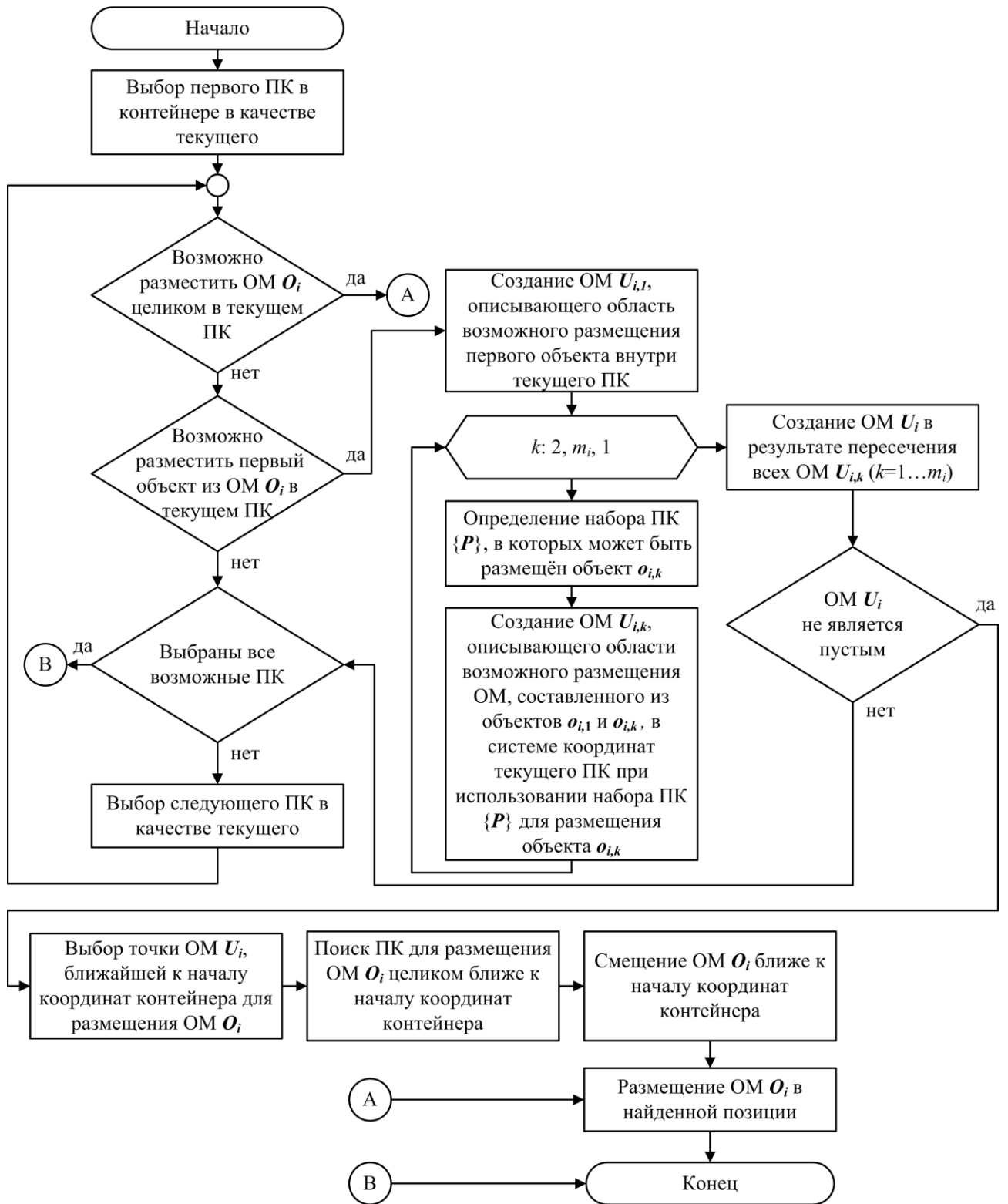
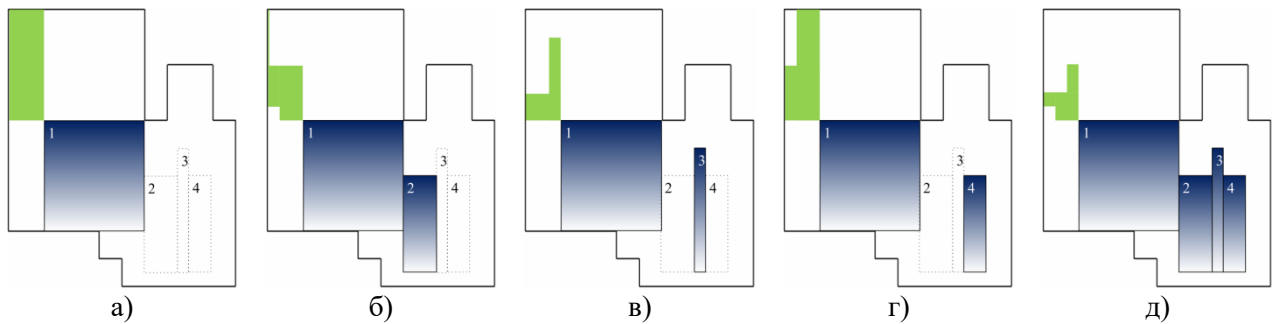


Рис. 2. Блок-схема алгоритма размещения ортогонального многогранника в контейнере

Процесс определения области возможного размещения ортогонального многогранника иллюстрирует пример на рис. 3.



**Рис. 3. Определение области допустимого размещения двухмерного ортогонального многогранника (показан градиентной заливкой) при использовании ПК, расположенного в левом верхнем углу контейнера, для размещения первого объекта ОМ:
а – ОМ $U_{i,1}$; б – ОМ $U_{i,2}$; в – ОМ $U_{i,3}$; г – ОМ $U_{i,4}$; д – область допустимого размещения, полученная в результате применения операции пересечения к ОМ $U_{i,1}$, $U_{i,2}$, $U_{i,3}$ и $U_{i,4}$**

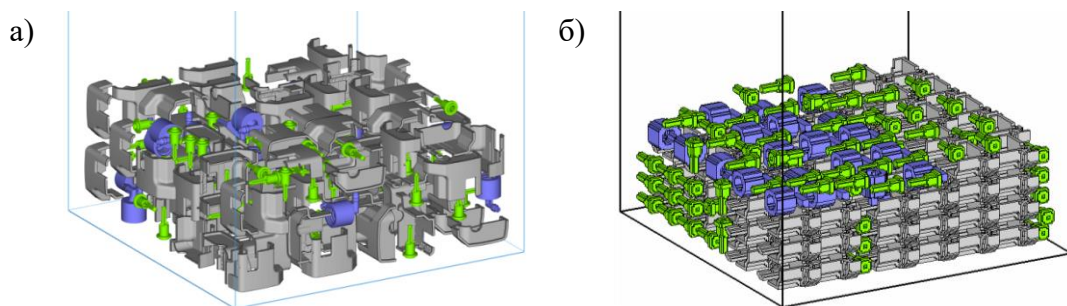
Решение практических задач компоновки

В настоящее время наиболее богатый функционал для решения задач планирования и подготовки аддитивного производства предоставляет программный продукт Materialise Magics, разработанный бельгийской компанией Materialise NV (<https://www.materialise.com/en/software/magics>), продукция которой используется 92% ведущих мировых производителей оборудования для трёхмерной печати (<https://www.materialise.com/en/facts>). Широкое признание программного продукта Materialise Magics на мировом рынке аддитивных технологий, а также наличие встроенного модуля Sinter, предназначенного для оптимизации размещения деталей на платформе, объясняет его выбор для проведения сравнительного анализа полученных результатов компоновки.

В рамках проведённых исследований были решены задачи компоновки деталей внутри трёхмерного контейнера в форме параллелепипеда длиной 340 мм, шириной 340 мм и высотой 620 мм. При формировании компоновок были разрешены повороты деталей на углы, кратные 90° , а также заданы следующие ограничения: интервал между размещаемыми деталями равен 6 мм, отступ от нижней грани контейнера равен 9 мм, отступ от боковых граней контейнера равен 10 мм.

Разработанные алгоритмы формирования и размещения ортогональных многогранников различной размерности реализованы в программном обеспечении Packer [16]. Вычислительные эксперименты проводились на персональном компьютере (ЦП – Intel Core i5-8400, 2,8 ГГц; 8 ГБ ОЗУ).

Результаты решения задачи компоновки 200 деталей трёх различных типов приведены на рис. 4.



**Рис. 4. Компоновка деталей на платформе 3D-принтера (задача 1):
а – наилучшее решение задачи, полученное модулем Sinter программы Materialise Magics (высота сборки: 141,0 мм, время решения: 300 с); б – компоновка, полученная разработанной программой Packer (высота сборки: 122,4 мм; время решения: 107 с)**

Результаты решения задачи компоновки 100 деталей восьми различных типов приведены на рис. 5.

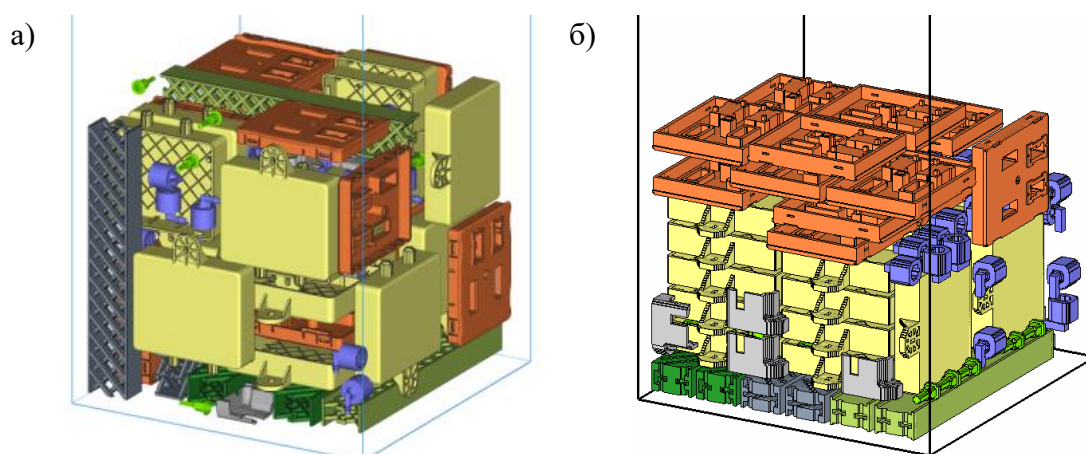


Рис. 5. Компоновка деталей на платформе 3D-принтера (задача 2):
а – наилучшее решение задачи, полученное модулем Sinter программы Materialise Magics (высота сборки: 318,0 мм, время решения: 300 с); б – компоновка, полученная разработанной программой Packer (высота сборки: 298,4 мм; время решения: 96 с)

Представленные примеры демонстрируют превосходство компоновок (как по скорости их формирования, так и по плотности размещения деталей), получаемых на основе разработанных алгоритмов формирования и размещения ортогональных многогранников, перед компоновками, формируемыми программой Materialise Magics.

Очевидно, что с увеличением размеров вокселей будет сильнее искажаться форма детали при вокселизации, а полученный ортогональный многогранник будет состоять из меньшего числа ортогональных объектов, что обеспечит рост скорости формирования компоновки. Например, сокращение числа ортогональных объектов в 2,5 раза для задачи 1 приводит к повышению скорости формирования компоновки в 6 раз, при этом плотность размещения объектов снижается всего на 2% (высота сборки: 124,4 мм; время решения: 18 с).

Выводы

Разработанный алгоритм размещения ортогональных многогранников решает задачу рациональной компоновки деталей сложной геометрической формы. Применение вокселизации для получения ортогональных многогранников позволяет управлять скоростью и точностью решения задачи компоновки. Важно отметить, что предложенный алгоритм реализован инвариантно относительно размерности размещаемых объектов, что делает его применимым также для задач фигурного раскроя листового материала.

Показано, что разработанные алгоритмы формирования и размещения ортогональных многогранников обеспечивают получение в среднем на 10% более плотных компоновок в сравнении с компоновками, формируемыми программным решением Materialise Magics, являющимся лидером на рынке программного обеспечения для аддитивного производства.

Библиографический список

1. Соболев А.Н. Моделирование конструкций корпусных деталей с использованием расчётных макроэлементов / А.Н. Соболев, М.Г. Косов, А.Я. Некрасов // Вестник МГТУ «Станкин». – 2014. – № 3 (30). – С. 98–101.
2. Чеканин В.А. Мультиметодный генетический алгоритм для решения задач раскроя и упаковки прямоугольных объектов / В.А. Чеканин // Вестник МГТУ «Станкин». 2019. № 4 (51). С. 14–18.

3. Чеканин В.А. Эвристический алгоритм оптимизации решений задачи прямоугольного раскроя / В.А. Чеканин, А.В. Чеканин // Вестник МГТУ «Станкин». – 2014. – № 4. – С. 210–213.
4. Математические модели и оптимизационные методы геометрического проектирования / Ю.Г. Стоян, С.В. Яковлев // Киев: Наук. думка, 1986. – 286 с.
5. Верхотуров М.А. Упаковка сложных трёхмерных объектов в прямоугольный контейнер на базе дискретно-логического представления информации / М.А. Верхотуров [и др.] // Известия Самарского научного центра Российской академии наук. – 2014. – Т. 16, № 4–2. – С. 378–383.
6. Ягудин Р.Р. Оптимизация компоновки трехмерных геометрических объектов на основе годографа вектор-функции плотного размещения / Р.Р. Ягудин // Инженерный вестник Дона. – 2012. – Т. 21, № 3. – С. 206–217.
7. Стоян Ю.Г. Моделирование плотной упаковки 3D-объектов / Ю.Г. Стоян, В.В. Сёмкин, А.М. Чугай // Кибернетика и системный анализ. – 2016. – С. 137–146.
8. Vucholm T. Effective packing of 3-dimensional voxel-based arbitrarily shaped particles / T. Vucholm, M. Toivakka, J. Westerholm // Powder Technology. – 2009. – Vol. 196, № 2. – P. 139–146.
9. Локтев М.А. Функциональный принцип обхода препятствий с применением метода функционально-воксельного моделирования / М.А. Локтев, А.В. Толок // Вестник МГТУ «Станкин». – 2016. – № 1 (36). – С. 75–80.
10. Тун Е.В. Дискретные модели обеспечения зон обслуживания и автоматизированной компоновки летательных аппаратов / Е.В. Тун, Л.В. Маркин // Труды МАИ. – 2017. – № 96. – С. 28.
11. Чеканин В.А. Алгоритмы формирования ортогональных многогранников произвольной размерности в задачах раскроя и упаковки / В.А. Чеканин, А.В. Чеканин // Вестник МГТУ «Станкин». – 2018. – № 3 (46). – С. 126–130.
12. Чеканин В.А. Алгоритмы размещения и удаления ортогональных многогранников произвольной размерности в задачах раскроя и упаковки / В.А. Чеканин // Вестник МГТУ «Станкин». – 2019. – № 3 (50). – С. 77–81.
13. Чеканин В.А. Алгоритм декомпозиции ортогонального многогранника произвольной размерности / В.А. Чеканин, А.В. Чеканин // Информационные системы и технологии - 2020: сборник материалов XXVI Международной научно-технической конференции – Н. Новгород: Нижегород. гос. техн. ун-т им. Р.Е. Алексеева, 2020. – С. 889–895.
14. Чеканин В.А. Структура данных для задачи трехмерной ортогональной упаковки объектов / В.А. Чеканин, А.В. Чеканин // Вестник МГТУ «Станкин». – 2015. – № 1. – С. 112–116.
15. Chekanin A.V. Effective data structure for the multidimensional orthogonal bin packing problems / A.V. Chekanin, V.A. Chekanin // Advanced Materials Research. – 2014. Vol. 962-965. P. 2868–2871.
16. Chekanin V.A. Design of Library of Metaheuristic Algorithms for Solving the Problems of Discrete Optimization / V.A. Chekanin, A.V. Chekanin // Advances in Mechanical Engineering. Lecture Notes in Mechanical Engineering. – Springer, Cham, 2018. – P. 25–32.

V.A. Chekanin

SOLVING THE PROBLEM OF RATIONAL LAYOUT OF PARTS OF COMPLEX GEOMETRIC SHAPE

Moscow State University of Technology «STANKIN»

Purpose: Development of an algorithm that provides fast solving the problem of dense layout of parts of complex geometric shape.

Approach: Initially, parts of a complex geometric shape are subjected to voxelization, which ensures their translation into a discrete representation, further the decomposition of the obtained orthogonal polyhedrons is performed to reduce the number of orthogonal objects constituting them, after which the developed algorithm for placing orthogonal polyhedron of arbitrary dimension is applied.

Findings: The high efficiency of algorithms developed for the formation and placement of orthogonal polyhedrons is demonstrated by examples of solving layout problems of parts on the platform of a 3D printer.

Value: The implemented approach to solving the problems of rational placement of parts of complex geometrical shapes provides fast formation of dense layouts. The developed algorithm for placing orthogonal polyhedrons selects the most suitable free areas of a container to obtain the most compact placement of all parts.

Keywords: layout problem, packing problem, orthogonal polyhedron, additive manufacturing.

ПРИМЕНЕНИЕ МОДИФИКАЦИИ МЕТОДА ВЕТВЕЙ И ГРАНИЦ ДЛЯ РЕШЕНИЯ ЗАДАЧИ КОММИВОЯЖЕРА

¹ НИУ Высшая школа экономики, г. Москва;

² Нижегородский государственный технический университет им. Р.Е. Алексеева

Классический алгоритм, реализующий метод ветвей и границ для решения задачи коммивояжера, предложенный в 1963 году Дж. Литл, К. Мурти, Д. Суини и К. Кэрл, остаётся по настоящее время одним из самых востребованных точных алгоритмов при решении задачи нахождения гамильтонова цикла минимальной стоимости в полном взвешенном неориентированном графе. Однако, экспоненциальная асимптотическая оценка сложности алгоритма не позволяет решать ряд прикладных задач большой размерности за время, приемлемое для лица, принимающего решение. Ввиду ограниченных аналитических способностей человека и сложной структуры задачи, анализ задачи коммивояжера и разработка ресурсно-эффективного алгоритма для её решения становится непростой задачей. В данной работе рассматриваются комбинации модификаций метода ветвей и границ, позволяющие сократить время решения задачи коммивояжера.

Ключевые слова: Задача коммивояжера, метод ветвей и границ, сложность индивидуальной задачи, оценка и анализ алгоритмов, поисковое дерево решений.

1. Введение

Целый ряд практических задач в области логистики сводится к классической задаче коммивояжера. Она заключается в поиске гамильтонова цикла с минимальной стоимостью в полном асимметричном графе. Одним из самых известных алгоритмов решения задачи коммивояжера является алгоритм метода ветвей и границ, предложенный Дж. Литл, К. Мурти, Д. Суини и К. Кэрл в 1963 году [1]. Однако, данный алгоритм имеет экспоненциальную временную сложность по размерности задачи (количество вершин в графе). Даже для не таких больших размерностей (уже при числе вершин графа порядка 30 – 50), время решения задачи коммивояжера является неприемлемым с точки зрения лица, принимающего решения. В различных работах по совершенствованию МВГ работами были предложены различные подходы, направленные на повышение временной эффективности метода ветвей и границ, например:

- использование дополнительной памяти [2];
- использование различных подходов к организации поискового дерева решений [3];
- использование предвычисленного тура, полученного с помощью различных метаэвристических алгоритмов, в качестве начального тура МВГ [4].

Ранее автором в соавторстве с другими исследователями был проведен анализ задачи коммивояжера и алгоритма её решения с помощью метода ветвей и границ, например:

- классификация задач по матрицам номеров порядка [5];
- оценка сложности индивидуальной задачи коммивояжера [6] согласно работе Дональда Кнута [7];
- вероятностный прогноз сложности индивидуальных задач коммивояжера на основе идентификации распределения сложности [8];
- введена новая оценка задачи для обоснования полученных результатов [9].

Более того, за счёт искусственного ограничения числа вершин поискового дерева решений, которое строится в ходе работы метода ветвей и границ, получилась

модификация алгоритма, которая не гарантирует оптимальный результат, но предоставляет решение, близкое к оптимальному [10].

Все упомянутые выше работы заложили фундамент для разработки модифицируемого алгоритма метода ветвей и границ, который решает задачу коммивояжера, являющейся подзадачей в рамках задачи построения маршрутов и графика грузовых перевозок автодорожным транспортом. Об этой модификации алгоритма и пойдет речь в данной статье.

2. Основание для разработки

Задача построения маршрутов и графика грузовых перевозок автодорожным транспортом возникла в частной транспортной компании, которая специализируется на автоперевозках готовых автомобилей, контейнеров и генеральных грузов.

Для подбора маршрутов и рассмотрения возможностей смешанных перевозок грузов компания разработала собственный алгоритм, одним из этапов которого является решение пакета классических задач коммивояжера. На данный момент, для решения классической задачи коммивояжера используется жадный алгоритм. Сотрудники компании понимают, что решения, полученные жадным алгоритмом, далеки от оптимального, однако, алгоритм работает очень быстро и прост в реализации. Также, по оценкам сотрудников, решение задачи коммивояжера с помощью классической реализации метода ветвей и границ существенно увеличит время работы алгоритма построения всех маршрутов, что не допустимо с точки зрения бизнес требований. Отметим, что размерность задач варьируется в интервале от 30 до 50. Очевидно, что получение точных решений для возникающего пакета классических задач коммивояжера позволит получить более экономически эффективное расписание, но со стороны разработчиков выдвигаются существенные требования по временной эффективности.

3. Метод ветвей и границ

В данной работе не будет уделяться внимание всем аспектам метода ветвей и границ для решения задачи коммивояжера, однако, необходимо уделить внимание тем этапам алгоритма, которые имеют непосредственное значение для разрабатываемой модификации алгоритма. За деталями реализации классического метода ветвей и границ для решения задачи коммивояжера можно обратиться к следующим литературным источникам: [1], [11], [12]. Ниже представлен псевдокод алгоритма метода ветвей и границ для задачи коммивояжера:

A(C, n) псевдокод алгоритма метода ветвей и границ для задачи коммивояжера

(n – размерность, C – матрица стоимостей.)

1. Инициализация.

2. Приведение матрицы стоимостей C.

3. Установка корня поискового дерева решений $X=R$
приведение исходной матрицы – вычисление $w(X)$.

While ($w(X) < z_0$)

begin

4. Выбор ребра ветвления (k, l).

5. Процесс ветвления. Создание вершины Y^-
и вычисление $w(Y^-)$.

6. Процесс ветвления. Создание вершины Y
и вычисление $w(Y)$.

If (размер матрицы стоимостей в вершине $Y = 2$)

then

begin

7. Проведение исчерпывающей оценки для вершины Y

If ($w(Y) < z_0$)

then

```

begin
    z0 = w(Y) (запоминаем тур)
end
end
8. Выбор следующей вершины поискового дерева
решений, и установка X
9. Вычисление фрагмента матрицы C,
соответствующего выбранной вершине X,
на основании пути от корня поискового
дерева решений до текущей вершины.
end (while w(X) < z0)
Оптимальное решение со стоимостью z0 найдено
End.

```

Рассмотрим некоторые этапы и важные детали этого алгоритма:

- *Этап 5.* На этом этапе реализуется ветвление, и в поисковое дерево добавляется вершина Y^- , следующая за X .
- *Этап 6.* Формирование вершины Y , следующей за X , соответствует подмножество туров из множества X , содержащих то ребро (k, l) , которое выбрано на этапе 4.
- *Этап 7.* В конечном итоге процесс ветвления приводит нас к множествам, содержащим так мало туров, что можно рассмотреть каждый из них и провести оценку для этой вершины без дальнейшего ветвления. После проведения оценки необходимо запомнить тур как кандидата на оптимальный тур.
- *Этап 8.* На этом этапе необходимо выбрать следующую вершину X , от которой необходимо проводить ветвление. Этот выбор довольно очевиден: должна быть выбрана вершина поискового дерева решений, которая имеет в данный момент времени наименьшую нижнюю границу и из которой в данный момент времени не выходят ветви, (т.е. это лист поискового дерева решений с минимальной оценкой).

4. Подход к разработке модификации алгоритма метода ветвей и границ

Для разработки эффективной по времени модификации метода ветвей и границ для решения задачи коммивояжера необходимо провести экспериментальное исследование на основе переданных от компании матриц стоимостей и определить, какие модификации сократят время работы классического алгоритма метода ветвей и границ.

Необходимо рассмотреть применение как отдельных, так и совокупности следующих подходов к повышению временной эффективности:

- использование дополнительной памяти;
- использование различных подходов к организации поискового дерева решений;
- использование в качестве начального тура предвычисленный тур.

Такая совокупная модификация должна позволить существенно сократить время работы классического алгоритма метода ветвей и границ, и, в тоже время, предоставить точное решение задачи коммивояжера. Полученное на этой основе программное решение должно позволить компании, без ущерба их основному бизнесу, получить в динамике заказов более эффективное расписание, что позволит существенно сократить издержки и увеличить объём заказов (без увеличения автопарка).

Библиографический список

1. Little J. D. C., Murty K. G., D.W. Sweeney, and Karel C., "An algorithm for the traveling salesman problem," Operations Research, vol. 11, pp. 972–989, 1963.
2. Ulyanov M. V., Fomichev M. I., Resource characteristics of ways to organize a decision tree in the branch-and-bound method for the traveling salesman problem, Business Informatics, no. 34, pp. 38-46, Dec. 2015.

3. Ульянов М. В., Фомичев М.И., Подходы к организации поискового дерева решений в методе ветвей и границ для асимметричной задачи коммивояжера // Информационные технологии. 2018. Т. 24. № 11. С. 698-704.
4. Ульянов М. В., Фомичев М. И. Сравнительный анализ комбинаций метода ветвей и границ с метаэвристическими алгоритмами для решения асимметричной задачи коммивояжера // Информационные технологии. 2019. Т. 25. № 10. С. 590-595.
5. Жукова Г. Н., Ульянов М. В., Фомичев М. И., Головешкин В. А. Об одном обобщённом представлении классов индивидуальных задач коммивояжера // Автоматизация. Современные технологии. 2016. № 10. С. 22-29.
6. Головешкин В. А., Жукова Г. Н., Ульянов М. В., Фомичев М. И. Корреляция сложности и времени решения TSP // Системы компьютерной математики и их приложения. 2017. № 18. С. 136-138.
7. Knuth, D. E., Estimating the Efficiency of backtracking programs. *Mathematics of Computing*, v. 29, pp. 121–136, 1975.
8. Ульянов М. В., Жукова Г. Н., Фомичев М. И., Головешкин В. А. Вероятностный прогноз сложности индивидуальных задач коммивояжера на основе идентификации распределения сложности по экспериментальным данным // Автоматика и телемеханика // Автоматика и телемеханика. 2018. № 7. С. 149-166.
9. Фомичев М. И. Об одной оценке индивидуальной задачи коммивояжера // Информационные системы и технологии – 2019. Сборник материалов XXV Международной научно-технической конференции. 2019. С. 694-698.
10. Фомичев М. И. Об искусственном ограничении числа вершин поискового дерева решений для решения задачи коммивояжера методом ветвей и границ // Информационные системы и технологии – 2020. Сборник материалов XXVI Международной научно-технической конференции. 2020. С. 804-808.
11. Гудман С., Хидетниemi С. Ведение в разработку и анализ алгоритмов. / – М.: Мир, 1981. – 368 с.
12. Ульянов М.В. Ресурсно-эффективные компьютерные алгоритмы. Разработка и анализ. / М.: ФИЗМАТЛИТ, 2008 – 304 с. ISBN 978-5-9221-0950-5.

M. I. Fomichev^{1,2}

APPLYING OF BRANCH AND BOUND METHOD MODIFICATION TO SOLVE THE TRAVELLING SALESMAN PROBLEM

¹*NRU HSE, Moscow;*

²*Nizhny Novgorod State Technical University n.a. R.E. Alekseev*

The classical Branch and Bound algorithm for solving the Traveling Salesman Problem, introduced in 1963 by J. D. C. Little, K. G. Murty, D. W. Sweeney, and C. Karel, is still one of the most applied accurate algorithms to find Minimum Cost Hamiltonian Circuit in a complete weighted undirected graph. However, exponential asymptotic computational complexity estimation of the algorithm does not allow to solve a number of applied large dimension problems for time, which is acceptable to the decision-maker. Due to limited analytical person abilities and complex problem structure, the analysis of the Travelling Salesman Problem and the development of a resource-efficient algorithm to solve it becomes an uneasy task. This paper presents combinations of different modifications of the Branch and Bound Method to reduce time for resolving Traveling Salesman Problem.

Keywords: Travelling Salesman Problem, Branch and Bound Method, complexity of the individual problem, evaluation and analysis of algorithms, search decision tree.

П.А. Козлов, Н.А. Назаров, В.О. Толчеев

ПЕРСОНАЛИЗИРОВАННАЯ СИСТЕМА ПОДДЕРЖКИ НАУЧНОЙ ДЕЯТЕЛЬНОСТИ В МАЛЫХ ИССЛЕДОВАТЕЛЬСКИХ КОЛЛЕКТИВАХ

Национальный исследовательский университет «Московский энергетический институт»

В работе представляется персонализированная система поддержки научной деятельности, предназначенная для работы в малых исследовательских коллективах. Апробация системы проводится на примере кафедры Управления и интеллектуальных технологий НИУ «МЭИ». В настоящее время реализовано два модуля, которые позволяют давать рекомендации по выбору профильных научных конференций и осуществлять фильтрацию научных публикаций согласно профессиональным интересам сотрудников.

Ключевые слова: интеллектуальный анализ данных, косинусная мера, бинарная классификация, обработка текста, рекомендательная система, построение профилей, ROC-кривая.

Введение

В настоящее время информационная потребность специалистов, работающих в научно-образовательной сфере, заключается в своевременном получении сведений, полезных для проведения собственных исследований. В частности, актуальным представляется информирование научных сотрудников и преподавателей о проводимых профильных научных конференциях и издаваемых релевантных публикациях.

Для облегчения доступа пользователей (ученых, специалистов, преподавателей, аспирантов, студентов) к научным информационным ресурсам разработаны специальные информационно-поисковые системы (ИПС) Web of Science, Scopus, Microsoft Academic, Google Scholar, Cite Seer, Sci-hub и т.п. Большинство таких ИПС осуществляет поиск в англоязычных базах данных и не охватывает русскоязычный массив специализированной информации, представленный, в частности, в научной электронной библиотеке e-Library. Кроме того, известные ИПС нацелены прежде всего на выявление релевантной информации, которая будет лишь частично интересна пользователю. Только некоторые программы имеют встроенные функции для аналитики и визуализации результатов (например, VOS Viewer для Web of Science) [1].

В связи с этим представляется актуальным разработать персонализированную систему поддержки научной деятельности, которая в автоматизированном режиме будет предоставлять специалистам востребованные (по возможности пертинентные) сведения о русскоязычных конференциях и научных публикациях, соответствующих их профессиональным интересам. В отличие от общеизвестных ИПС, наша разработка ориентирована на малые научные коллективы (кафедра, лаборатория, отдел). В этом случае можно улучшить качество поиска, обработки и анализа текстовых данных за счет использования имеющейся персональной информации о сотрудниках.

Далее в качестве потенциальных пользователей разрабатываемой системы рассматриваются преподаватели *кафедры управления и интеллектуальных технологий НИУ МЭИ*. Данная кафедра осуществляет подготовку студентов (и НИР) по дисциплинам Информатики (Computer Science). Вместе с тем часть преподавателей проводит исследования в области Интеллектуального анализа данных (ИАД, Data and Text Mining). Рассматриваемая в работе персонализированная система ориентирована на специалистов в ИАД и способна на основе их предпочтений предоставлять актуальную профильную информацию о предстоящих научных конференциях и отслеживать публикации по тематике ИАД в документальном потоке по Информатике.

Предварительная обработка текстовых данных и постановка задачи

Для разработки персонализированной системы поддержки научной деятельности нами сформирована выборка из цифровой библиотеки e-Library, состоящая из библиографических описаний статей (авторы, название, аннотация, ключевые слова).

Для описания документа используется векторная модель X_j :

$$X_j = \begin{bmatrix} x_{1j} \\ x_{ij} \\ x_{Nj} \end{bmatrix}, (1)$$

где x_{ij} – вес термина i в документе j

($j = 1, \dots, M$ – количество документов в выборке, $i = 1, \dots, N$ – количество признаков)

Для описания выборки использовалась матричная модель X :

$$X = \begin{pmatrix} x_{11} & \dots & x_{1M} \\ \dots & x_{ij} & \dots \\ x_{N1} & \dots & x_{NM} \end{pmatrix} (2)$$

Такую матрицу называют матрицей «термин - документ», т.к. её строки представляют собой документы, а столбцы – термины, содержащиеся в этих документах.

Эта матрица получается после проведения предварительной обработки текстовых документов, которая включает:

- 1) Очистка слов от знаков препинания и ненужных символов.
- 2) Приведение текста к одному регистру.
- 3) Разбиение текста на отдельные слова.
- 4) Лемматизация
- 5) Удаление стоп-слов.
- 6) Удаление низкочастотных слов

Кроме того, составлялись профили T_j и Z_g – векторные описания пользователей и научных конференций. Профили строятся из библиографических описаний статей преподавателя.

$$T_j = \begin{bmatrix} t_{1j} \\ t_{ij} \\ t_{nj} \end{bmatrix}, (3)$$

где T_j – профиль j -го сотрудника, t_{ij} – вес термина i в профиле j ($i = 1, \dots, n$ – количество признаков, $j = 1, \dots, m$ – количество сотрудников).

$$Z_j = \begin{bmatrix} z_{1j} \\ z_{ij} \\ z_{nj} \end{bmatrix}, (4)$$

где Z_j – профиль j -ой конференции, z_{ij} – вес термина i в описании конференции j

($j = 1, \dots, Q$ – количество конференций)

Дополнительно из публикаций специалистов извлекаются сведения о конференциях, в которых они участвовали (название конференции и её описание).

Для расчета весов x_{ij} , t_{ij} и z_{ij} используется tfс – взвешивание, т.к. оно усредняет частоты слов, занижая частоты высокочастотных слов и увеличивая частоту низкочастотных слов [5].

$$x_{ij} = \frac{f_{ij} * \log\left(\frac{M}{M_i}\right)}{\sqrt{\sum_{i=1}^N [f_{ij} * \log\left(\frac{M}{M_i}\right)]^2}}, (5)$$

где f_{ij} – частота встречаемости i -го слова в j -ом документе, а M_i – количество документов, в которых встретилось i -ое слово. Суммирование в знаменателе дроби проводится по всем терминам j -го документа, в котором встречается i -ое слово.

Аналогично с использованием t_{fc} – взвешивания рассчитываются веса терминов t_{ij} из z_{ij} в профилях сотрудников и конференций.

Сделаем формальную постановку задачи:

На основе сформированной выборки и построенных профилей в нашей работе решаются следующие задачи:

- выявляются релевантные конференции для рекомендации преподавателям,
- обучается бинарный классификатор для выявления релевантных публикаций для просмотра преподавателями.

Таким образом, разрабатываемая персонализированная система поддержки научной деятельности состоит из нескольких модулей: мониторинга специализированных информационных ресурсов, предварительной обработки текстовых данных, выявления профильных конференций и фильтрации (бинарной классификации) публикаций с учетом профессиональных интересов пользователя.

Разработка рекомендательной системы по выбору профильных научных конференций

В модуле выбора профильных научных конференций используется фильтрация на основании содержимого (Content-based filtering) [2]. Этот подход эффективно применяется при разработке рекомендательных систем и предполагает, что мы имеем достаточно информации о пользователях и рекомендуемых им объектах (конференции). Необходимо определить предпочтения пользователя и на основании этого выдать ему рекомендации.

В ходе разработки системы было принято решение использовать всю имеющуюся информацию о пользователях – строить профили преподавателей на основе их публикаций (а не только докладов на конференциях).

Измерять удовлетворенность пользователя предложенной конференцией будем, используя косинусную меру, значения которой могут изменяться на отрезке $[0;1]$ [3].

$$h_{j,q} = \frac{(T_j, Z_q)}{|T_j| * |Z_q|} = \frac{\sum_{i=1}^n t_{ij} * z_{iq}}{\sqrt{\sum_{i=1}^n t_{ij}^2} * \sqrt{\sum_{i=1}^n z_{iq}^2}}, \quad (6)$$

где $H_{j, q}$ – значение косинусной меры между j -ым сотрудником и q -ой конференцией; Z_q – вектор весов слов q -ой конференции; T_j – вектор весов слов профиля j -ого сотрудника.

Для формирования множества Z написано 3 программы-парсера для сайтов, содержащих информацию о конференциях:

- 1) <https://www.science-community.org/ru/>;
- 2) <https://проконференции.рф/>;
- 3) <https://konferencii.ru/>.

Эти сайты выбраны так как они постоянно добавляют актуальную информацию о новых конференциях и позволяют просмотреть основные данные обоих характеристиках (научные направления, секции, организаторы).

Так как мы используем несколько источников информации о новых конференциях, существует проблема дубликатов, для решения этой проблемы был реализован алгоритм по выявлению не уникальных конференций. Для поиска дубликатов был выбран метод Шинглов так как он является стандартным для поиска нечетких дубликатов [4].

Качество работы системы будем оценивать с помощью меры «точность», рассчитываемой по формуле [3]:

$$Precision = \frac{TP}{FP + TP}, \quad (7)$$

где TP - истинно-положительное решение; FP - ложно-положительное решение.

Использование точности связано с невозможностью расчета полноты для распространенной меры «полнота-точность» так как отсутствует размеченная выборка. Ее формирование вручную не представляется возможным в связи с высокой трудоемкостью.

Анализ результатов работы контентно-ориентированной процедуры показывает, что системе требуются доработки, улучшающие качество выдаваемых рекомендаций (точность оказалась 75%). С целью повышения качества работы системы реализована обратная связь по релевантности, в рамках которой пользователю предоставляется возможность давать оценку релевантности полученных рекомендаций по пятибалльной шкале. На основе оценок пользователя модифицируется его профиль путем увеличения весов у наиболее важных терминов.

После внедрения обратной связи по релевантности точность достигает 94%, что указывает на то, что данный алгоритм оказывает существенное влияние на качество работы системы и способен улучшить полезность рекомендаций.

Разработка системы фильтрации профильных научных публикаций

В модуле выбора профильных научных публикаций формировалась выборка, состоящая из библиографических описаний статей, хранящихся в цифровой библиотеке e-Library.ru и соответствующих запросу «Интеллектуальный анализ данных» (метка класса «ИАД»).

Фильтрация статей будет производиться путём бинарной классификации. Следовательно, после формирования тематики «ИАД» требуется сформировать тематику «не ИАД». Стоит заметить, что наборы статей по тематикам должны быть примерно одинакового размера, для лучшего обучения классификаторов. Так же отметим, что брать любые статьи будет нецелесообразно, т.к. в основном требуется разделить статьи, относящиеся к общей тематике «Computer Science» (широкая область специализации кафедры УИТ).

Статьи в класс «не ИАД» отбираются из журналов, в которых часто издаются работы по «ИАД», а также путем случайного отбора публикаций по запросу «Информатика» в e-Library (из статей до 2019 года включительно).

Размер полученной выборки равен 3100 статей, где 1600 статей относятся к тематике «ИАД» и 1400 статей к тематике «не ИАД». Данная выборка является обучающей, настройка параметров проводится с помощью пятикратной кросс-валидации. Кроме того, сформирована выборка из публикаций за 2020 год, состоящая из 400 статей, где 200 статей относятся к тематике «ИАД» и 200 статей к тематике «не ИАД». Сформированная выборка за 2020 год является тестовой выборкой.

Качество работы системы будем оценивать с помощью меры «Accuracy», рассчитываемой по формуле:

$$Accuracy = \frac{TP + TN}{TP + FP + TN + FN}, \quad (8)$$

где TP и FP поясняются в формуле (7), FN — ложно-отрицательное решение, TN — истинно-отрицательное решение.

Для проведения бинарной классификации выбраны классификаторы «Дерево решений», «Случайный лес», «Наивный Байес», «Логистическая регрессия», «Метод опорных векторов», «K-ближайших соседей» [6], реализованные в python библиотеки scikit-learn [7]. Параметры алгоритмов настроены по методу сетки, тестирование проводится на публикациях, опубликованных в 2020 году.

Выбор классификатора для фильтрации статей производится на основе построения и анализа ROC – кривых, расчёта значений AUC и Accuracy для тестовой выборки. ROC-кривая – это кривая показывающая зависимость количества верно классифицированных положительных объектов (чувствительность) от количества неверно классифицированных отрицательных объектов (специфичность). Мера AUC рассчитывается как площадь под ROC – кривой. Полученные результаты представлены на рисунке 1.

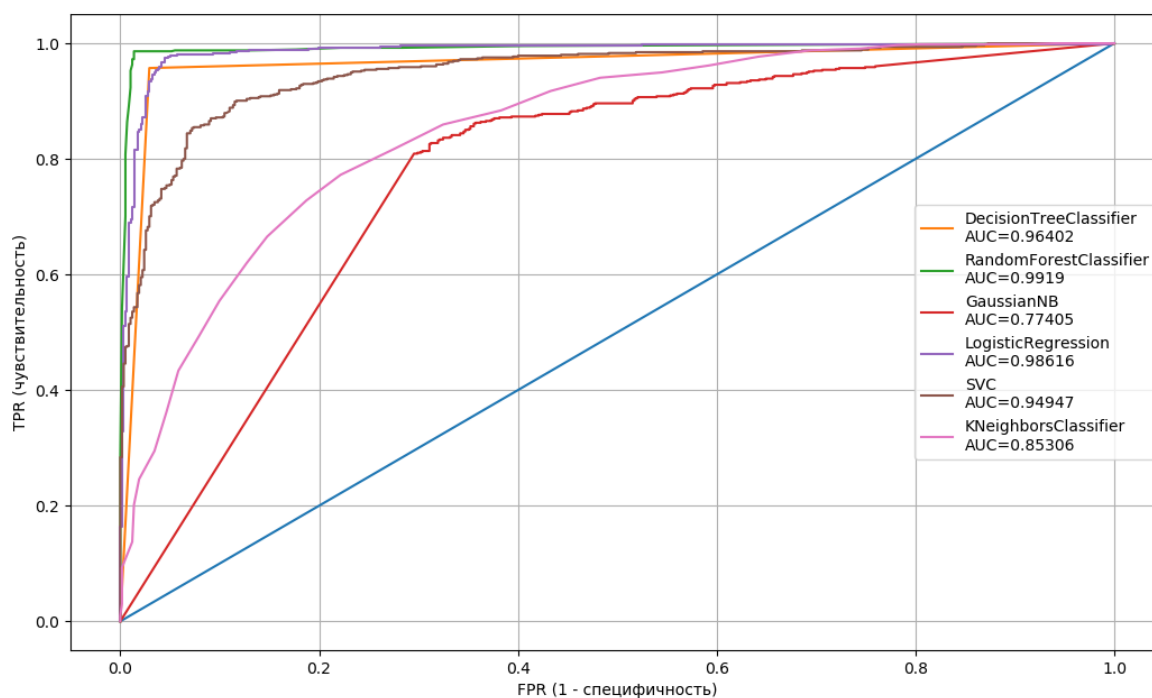


Рис. 1. ROC – кривые и AUC значения классификаторов тестовой выборки.

В качестве бинарного классификатора используется «Случайный лес». Данный классификатор считается одним из самых лучших алгоритмов машинного обучения, при обучении значение *Accuracy* достигло 0.99, на рисунке 1 ROC – кривая данного классификатора ближе всех находится к точке (0, 1). На тестовой выборке «Случайный лес» показал значение *Accuracy* равное 0.97.

Выводы

В ходе выполнения работы была разработана персонализированная система поддержки научной деятельности. Система апробирована на кафедре УИТ для сотрудников, специализирующихся в области Интеллектуального анализа данных. Результатом ее работы является список названий рекомендованных конференций с учетом профилей преподавателей, а также список научных публикаций по ИАД, рекомендованных к прочтению специалистам-предметникам.

Анализ работы системы показал, что она достаточно точно определяет специализированные конференции. В дальнейшем для улучшения точности планируется учесть дополнительные (зачастую субъективные) факторы, влияющие на выбор пользователя (место и время проведения, форма проведения, организаторы, индексация в Scopus и Web of Science и т.п.).

Наши эксперименты позволяют выбрать «Случайный лес» как наиболее точный метод фильтрации документов и выявления публикаций, которые соответствуют информационной потребности специалистов в области ИАД. Для достижения пертинентности результатов планируется реализовать дополнительные алгоритмы выявления информативных терминов и при необходимости проводить многоклассовую классификацию.

Библиографический список

1. Van Eck N. J., Waltman L. Software survey: VOS viewer, a computer program for bibliometric mapping //scientometrics. 2010. Vol. 84. №. 2. P. 523-538.
2. Koren Y., Bell R.M., Volinsky C., Matrix Factorization Techniques for Recommender Systems. IEEE Computer 42(8), 2009, pp. 42-49
3. Маннинг К., Рагхаван П., Шютце Х. Введение в информационный поиск. – М.: Вильямс, 2014 –528 с.
4. Цимбалов А.В., Золотарев О.В. Метод шинглов. Вестник российского нового университета. Серия: сложные системы, модели, анализ и управление. - 2016. - №4. - С. 72-79.
5. Borodkin A., Lisin E., Strielkowski V. Data algorithms for processing and analysis of unstructured text documents Applied Mathematical Sciences, Vol. 8, 2014, no. 25, 1213-1222.
6. Флах П. Машинное обучение – Наука и искусство построения алгоритмов, которые извлекают знания из данных. М.: ДМК-пресс. 2015. – 400с.
7. Библиотека scikit-learn. Machine Learning in Python: <https://www.sklearn.org>

P.A. Kozlov, N.A. Nazarov, V.O. Tolcheev

PERSONALIZED SUPPORT SYSTEM FOR SMALL RESEARCH TEAMS

National Research University “Moscow Power Engineering Institute”

The paper develops a personalized system for supporting scientific activities in small research teams. Testing of the system is carried out on the example of the Department of Control and Intelligent Technologies of Moscow Power Engineering Institute. Currently, two modules are implemented that allow to make recommendations on the selection of specialized scientific conferences and filter scientific publications according to the professional interests of user (specialist, tutor, student).

Keywords: data mining, cosine measure, binary classification, text processing, recommendation system, building profiles, ROC curve.

С.Р. Маслихов, А.С. Мохов, В.О. Толчеев

ИССЛЕДОВАНИЕ МЕТОДОВ СЕМПЛИНГА В ЗАВИСИМОСТИ ОТ УРОВНЯ НЕСБАЛАНСИРОВАННОСТИ КЛАССОВ В ЗАДАЧЕ КЛАССИФИКАЦИИ НОВОСТНЫХ СТАТЕЙ

Национальный исследовательский университет «МЭИ», г. Москва

Данная статья посвящена изучению влияния уровня несбалансированности классов на прирост качества классификации, измеряемое F1-мерой, получаемое после применения методов семплинга. В работе проводится сравнительный анализ двух основных подходов – методов уменьшения размера мажоритарных классов (undersampling) и методов увеличения размера миноритарных классов (oversampling). Исследования проводятся на выборке, сформированной из новостной коллекции Reuters. Даны рекомендации по применению рассмотренных методов.

Ключевые слова: интеллектуальный анализ данных, классификация текстов, новостные статьи, сэмплирование данных, undersampling, oversampling, несбалансированные выборки.

Введение

С каждым годом прогресс человечества в технической сфере открывает доступ к все более мощным вычислительным ресурсам. Этот факт в значительной степени влияет на развитие науки о данных в целом, и ее раздела, машинного обучения, в частности. Одной из задач машинного обучения является обучение с учителем (supervised learning), которая состоит в сопоставлении входным данным определенных меток. Для ее решения мы имеем обучающую выборку, в которой каждый объект характеризуется его признаковым описанием и меткой класса. Нередко приходится сталкиваться с неравномерным распределением объемов классов – так называемым дисбалансом классов. Следствием неравномерного обучения модели (классификатора) на разных по объему классах является как неверное распознавание объектов малых классов, так и трудности в оценке ее качества. По этим причинам необходимо искать подходы к выравниванию распределения объектов обучающей выборки.

Для борьбы с дисбалансом существует ряд методов семплинга (сэмплирования), разделяющихся на два вида: oversampling (генерация дополнительных объектов меньшего (миноритарного) класса) и undersampling (удаление объектов большего (мажоритарного) класса).

В данной работе исследуются методы семплинга для решения задачи классификации текстовых документов, т.е. требуется обучить классификатор относить новый документ к одному из заранее заданных классов. Исследуется целесообразность применения тех или иных методов семплинга в случае разного дисбаланса классов в выборке.

Решение задачи делится на 3 этапа:

1. Предварительная обработка данных.
2. Формирование базовой обучающей выборки.
3. Введение искусственного дисбаланса и проведение исследований.

Формирование выборки

В качестве выборки используется выборка Reuters, которая представляет собой коллекцию новостных статей, полученную из библиотеки nltk (язык программирования Python) и разделенную на несколько категорий (классов).

Изначально выборка имеет существенный дисбаланс. Объем выборки составляет 10,788 объектов, количество классов - 90. Каждый текстовый документ имеет множество меток, поэтому первым шагом в обработке выборки является сведение множества для каждого объекта к одной метке.

Подсчитаем число меток для каждой категории в выборке, возьмем первые 5 категорий, которые имеют не менее 500 документов (таблица 1, столбец №2). Сформируем базовую обучающую выборку - сбалансированную выборку, имеющую по 500 объектов в каждом классе, каждый объект которой принадлежит только к одному классу. Для каждого документа, из множества классов, к которым он относится, будем брать тот, объем которого наименьший (столбец №2). В столбце №3 показан состав выборки после приведения каждого документа к одному классу. Оставшиеся объекты использовались в качестве тестовой выборки.

Таблица 1

Обработка выборки

Категория	Объем с пересекающимися классами	Финальный объем
Earn (заработок)	3964	3931
Acq (acquisitions, приобретения)	2369	2345
Money-FX (деньги и биржа)	717	713
Grain (зерно)	582	580
Crude (сырье)	578	578

Важным шагом в решении задачи классификации является визуализация выборки. Эта проблема решается наиболее эффективно с использованием метода t-SNE [4]:

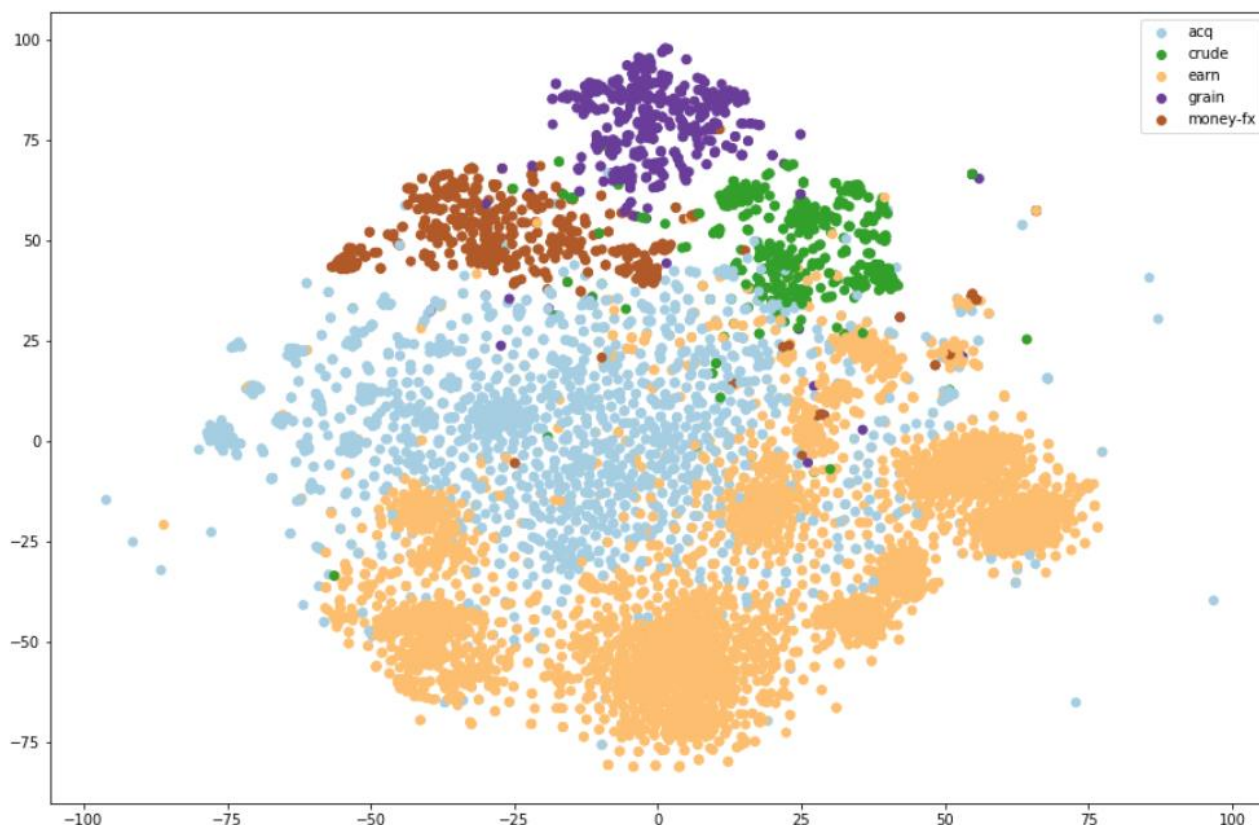


Рис. 1. Кластерная структура выборки

необходимо переобучить классификатор, тогда число классификаторов для каждого порога будет $n*k$, где k – число методов семплирования.

Чтобы иметь возможность проинтерпретировать результаты, для каждого метода мы будем усреднять показатели качества по разбиениям для каждого порога.

В качестве показателя качества будет использоваться F1-мера с микроусреднением по классам [3].

$$precision = \frac{TP}{TP+FP} \quad (2)$$

$$recall = \frac{TP}{TP+FN}, \quad (3)$$

где TP (True Positive), FP (False Positive), FN (False Negative) – показатели матрицы ошибок (confusion matrix) [2].

$$F1 = \frac{2*precision*recall}{precision+recall}, \quad (4)$$

Выбор микроусреднения оправдан тем, что в этом случае при расчете F1-меры не учитывается принадлежность объекта классу, т.е. каждый объект вносит одинаковый вклад в расчет меры. Общие значения TP, FP, FN, TN рассчитываются как сумма составляющих по каждому классу, и затем по общим значениям рассчитывается сама F1-мера.

Результаты классификации приведены на рисунке 3.

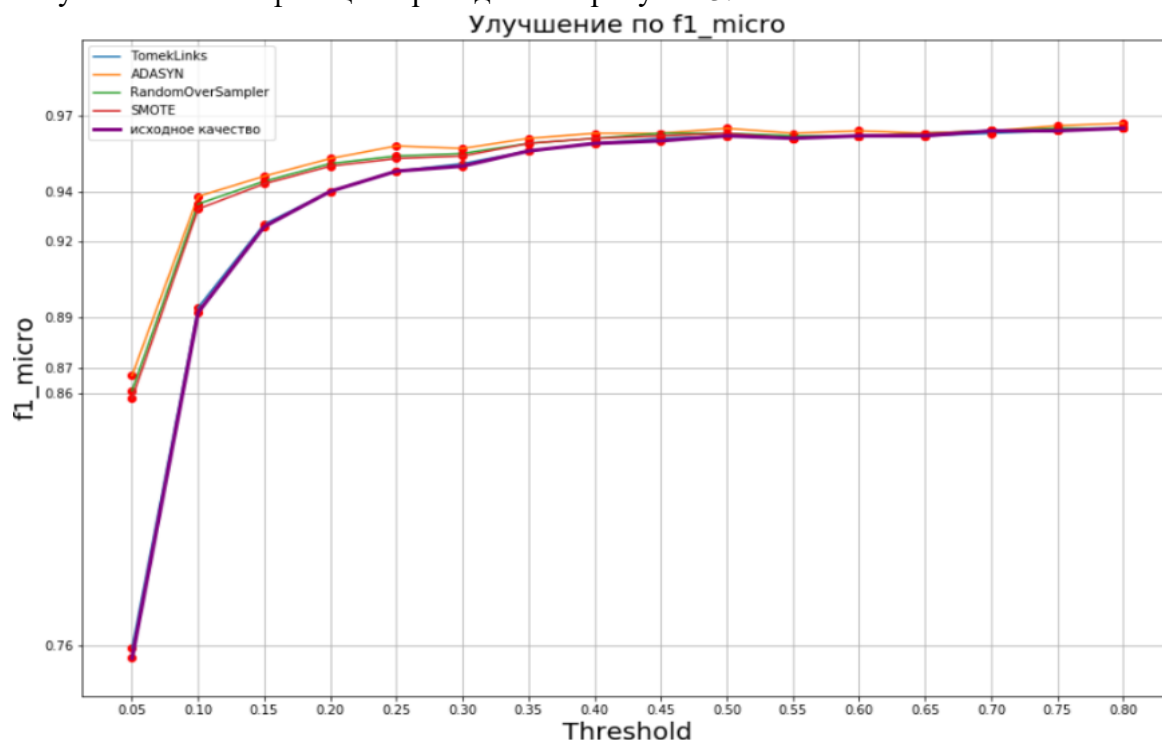


Рис.3. Среднее качество классификации для разных методов сэмпинга

Из рисунка видно, что методы oversampling показывают приблизительно одинаковый результат, при этом лучшим является ADASYN. Данный вывод подтверждается и исходной статьей, в которой метод ADASYN был представлен [6].

В таблице 3 представлены значения F1-micro для исходной выборки и метода ADASYN, а также выигрыш качества

Таблица 2

Выигрыш качества ADASYN

Порог	F1- micro, исходная выборка	F1-micro, выборка с ADASYN	Выигрыш качества, %
0.05	0.755	0.867	11.0
0.10	0.892	0.938	4.6
0.15	0.926	0.946	2.0
0.20	0.940	0.953	1.3
0.25	0.948	0.957	0.9
0.30	0.950	0.957	0.7
0.35	0.956	0.961	0.5
0.40	0.959	0.963	0.4
0.45	0.960	0.962	0.2
0.50	0.962	0.965	0.3
0.55	0.961	0.963	0.2
0.60	0.962	0.964	0.2
0.65	0.962	0.964	0.2
0.70	0.964	0.965	0.1
0.75	0.964	0.966	0.2
0.80	0.965	0.966	0.1

Наибольший выигрыш от применения методов семплинга можно наблюдать при пороге 0.05 (каждый из миноритарных классов включает 5% от объема мажоритарного, т.е. 25 объектов при мажоритарном классе – 500 объектов), это можно объяснить тем, что на более маленьких порогах модель слабо обучается на исходных данных, поэтому наиболее остро встает вопрос об увеличении выборки, что делают oversampling методы. При порогах более 0.2 применение методов семплинга дает лишь незначительный прирост менее 1%, соответственно их применение более целесообразно при большем дисбалансе классов.

Метод TomekLinks, как представитель группы undersampling методов, практически не позволяет улучшить качество при таком объеме выборки. Данный метод целесообразно использовать для решения задач классификации в условиях зашумленности мажоритарных классов.

Выводы

Проведенные исследования показывают, что высокий уровень несбалансированности (значения порога менее 0.2) достаточно негативно сказывается на качестве классификации по причине недообученности модели на малых классах. Соответственно, при таких малых значениях порога целесообразно применять методы oversampling, которые показывают приблизительно одинаковую эффективность. При больших значениях порога, и, соответственно, при меньшем уровне дисбаланса, применение методов сэмпинга в наших исследованиях показали выигрыш не более 1%, и целесообразность их применения должна решаться отдельно для каждого конкретного случая.

В дальнейшем, аналогичные исследования планируется провести на других, более зашумленных выборках, а также при разных объемах Vol_max количестве классов и обобщить полученные выводы на разные исходные данные.

Библиографический список

1. [More A. Survey of resampling techniques for improving classification performance in unbalanced datasets. 2016.](https://arxiv.org/pdf/1608.06048.pdf) URL: <https://arxiv.org/pdf/1608.06048.pdf> (дата обращения 31.01.2021).
2. Маннинг К.Д., Рагхаван П., Шютце Х. «Введение в информационный поиск». – М.: «Вильямс», 2014.
3. Yangand Y., Liu X. “Are-examination of text categorization methods,” in 22nd Annual International ACM SIGIR Conference on Research and Development in Information Retrieval (SIGIR’99), pp. 42–49, 1999.
4. Maaten, L. v. d., & Hinton, G. Visualizing data using t-SNE. Journal of machine learning research, 9(Nov), pp. 2579-2605, 2008.
5. Маслихов С.Р., Мохов А.С., Толчеев В.О. - Построение сбалансированных классов в задаче классификации запросов пользователей / Дистанционные образовательные технологии: сборник трудов V Международной научно-практической конференции - Симферополь: ИТ «АРИАЛ», 2020. – с. 245 – 248.
6. He H., Bai Y., Garcia E. A. and Li S., “ADASYN: Adaptive synthetic sampling approach for imbalanced learning,” in Proc. IJCNN, Hong Kong, China, 2008, pp. 1322-1328.

S.R. Maslikhov, A.S. Mokhov, V.O. Tolcheev

THE RESEARCH OF SAMPLING ALGORITHMS EFFICIENCY DEPENDING ON CLASS IMBALANCE LEVEL IN CLASSIFICATION OF NEWS ARTICLES

National Research University “MPEI”, Moscow

This article is devoted to the study of the influence of the class imbalance level on the classification quality (using F1-measure) after the sampling methods applying. The paper presents a comparative analysis of two main approaches – methods for reducing the size of majority classes (undersampling) and methods for increasing the size of minority classes (oversampling). The research is conducted on a sample formed from the Reuters news collection. Recommendations on the application of the considered methods are given.

Keywords: text mining, text classification, news articles, oversampling, undersampling, imbalanced samples.

Д.И. Жилинский¹, В.Л. Елисеев^{1,2}

КОМПЛЕКС МОДЕЛИРОВАНИЯ АТАК НА ТИПОВОЕ УСТРОЙСТВО ИНТЕРНЕТА ВЕЩЕЙ

¹Национальный исследовательский университет «МЭИ»,

²АО «ИнфоТеКС»

Работа посвящена современным проблемам информационной безопасности умных устройств, подключенных к сети Интернет., Проведен анализ наиболее известных компьютерных эпидемий для таких устройств. Выявлены типовые виды атак и изложен подход к их выявлению. Разработан комплекс для моделирования атак с целью разработки и тестирования системы защиты.

Ключевые слова: сетевые атаки, устройства Интернета Вещей, комплекс моделирования.

Введение

В настоящее время значительное внимание привлекают вопросы, относящиеся к обеспечению информационной безопасности, в том числе, кибер-физических систем, к которым можно отнести системы промышленной автоматизации, интеллектуальные бытовые устройства и мобильные устройства персонального использования, называемые совокупно Интернетом Вещей.

В самом широком смысле термин Internet Of Things, сокращенно IoT (Интернет Вещей) - это некая концепция подключения любого устройства к Интернету и другим подключенным устройствам.

Сегодня из-за обилия «умных» устройств люди не придают особого значения их безопасности, вследствие чего становятся явной и легкой целью для атак злоумышленников, которые могут нанести огромный ущерб. Так же большинство пользователей далеки от мира информационных технологий и не имеют никаких представлений о том, как нужно эксплуатировать приборы и какие меры безопасности при этом соблюдать [1].

В связи с этим и происходит огромное количество попыток взлома и атак на технику, именуемую устройствам «Интернета Вещей». Пример – ботнет BetaBot. Количество атак, совершенных этим ботнетом, составляет более 13% общем объеме уникальных атак 2018 года. Зловред работал в 42 странах, наибольший интерес его создателей вызывали финансовые сервисы, социальные сети и крупные порталы [2].

Современные сетевые атаки представляют значительную угрозу, так как их цель заключается в проникновении в систему и её злонамеренной эксплуатации, а также похищение или шифрование данных. Такая ситуация обусловлена сочетанием многих факторов, среди которых как низкая квалификация пользователей, так и наличие ошибок программного обеспечения. Актуальной задачей является разработка подходов к предотвращению компьютерных эпидемий среди умных устройств. Один из таких подходов предполагает разработку детектора аномалий на умном устройстве. Для отработки подхода была разработана модель типового умного устройства [3][4]. В данной работе решается задача разработки стенда для моделирования атак на типовое умное устройство, а также имитация атак с инфицированного устройства на другие хосты.

Обзор литературы

Угрозы информационной безопасности могут привести к потере конфиденциальных данных и нарушению работы информационных и кибер-физических систем. Их можно классифицировать на два типа: естественные и искусственные. К естественным относятся природные явления, которые не зависят от человека, например, ураганы, наводнения, пожары и т.д. Искусственные угрозы зависят непосредственно от человека и могут быть преднамеренные и непреднамеренные. Непреднамеренные угрозы возникают из-за неосторожности, невнимательности и незнания. Примером таких угроз может быть установка программ, которые не входят в число необходимых для работы, в дальнейшем нарушающих работу системы, что и приводит к потере информации.

Обычно классификация атак ориентирована на отдельные аспекты моделирования атак. Такими аспектами могут быть формализмы представления данных и знаний об атаках, уровни модели OSI (Open Systems Interconnection) или модели DoD (Department of Defense), типы атак, возможности масштабирования, способность учитывать динамические характеристики атак (например, время, параллельные процессы, цепочки взаимосвязанных инцидентов) и т.д. С точки зрения формального представления данных и знаний, наиболее популярные типы моделей атак — это табличные (матричные) модели, логические модели, модели, основанные на графах, а также имитационные модели [5]. Методы имитационного моделирования атак наиболее интересны тем, что позволяют легко включать в состав стенда агенты, реализующие или детектирующие реальные атаки, сохраняя простоту имитационных моделей других агентов.

Основная часть

Основной задачей данной работы являлась разработка комплекса для моделирования типовых атак со стороны злоумышленника.

Комплекс предназначен для анализа защищенности автоматизированных систем (АС), в нашем случае макет системы климат контроля, путем имитации действий нарушителя. Входными данными для моделирования являются спецификации анализируемой АС, модель нарушителя и исходные показатели защищенности, включая интегральный показатель «уровень защищенности АС». Выходными данными является информация об анализируемой сети в виде журналов регистрации событий, отчетов и т.д. Для реализации были выбраны три типовые атаки:

- Атака грубой силы
- SQL-injection (SQL – инъекция)
- Shell-injection (или по-другому – Command-injection)

Brute-force атака (атака грубой силы) - это попытка взломать пароль пользователя или найти скрытую веб-страницу или найти ключ, используемый для шифрования сообщения, используя метод проб и ошибок и надеясь, в конце концов, угадать правильно. Нахождение пароля для конкретного пользователя или сайта может занять много времени, поэтому хакеры разработали инструменты для более быстрой работы.

SQL-инъекция – это тип инъекционной атаки, которая позволяет выполнять вредоносные операторы SQL. Злоумышленники могут использовать уязвимости SQL Injection, чтобы обойти меры безопасности приложения. Они могут обойти аутентификацию и авторизацию веб-страницы или веб-приложения и получить содержимое всей базы данных SQL. Они также могут использовать SQL-инъекцию для добавления, изменения и удаления записей в базе данных

Shell-injection (Command-injection) – это уязвимость устройств, использующих небезопасное обращение к терминалу или командной строке операционной системы, которая позволяет злоумышленнику выполнять произвольные команды операционной системы (ОС) на сервере, на котором выполняется приложение, и, как правило, полностью скомпрометировать приложение и все его данные.

Данные атаки можно классифицировать следующим образом:

Атака на проникновение – Brute-force. Злонамеренное получение персональных данных для доступа к системе с возможностью кражи персональных данных и нецелевого использования.

Атака для кражи данных – SQL-injection. Через уязвимые веб-формы умного устройства кража информации и личных данных пользователей системы.

Атака для эксплуатации – Shell-injection. Использование уязвимостей сервера для потенциальной возможности управления устройством как промежуточном звеном в ботнете, либо полностью вредоносным комплексом.

Основная идея проведения представленных атак заключается в следующем – во-первых, возможность формирования обучающей выборки для детектора аномалий устройства Интернета Вещей, а во-вторых, проверка корректности работы данного детектора, насколько точно он будет реагировать на изменение сетевого трафика при попытке провести на него атаку.

В основе комплекса лежит виртуальная машина с оперативной Unix- системой Kali Linux. Kali Linux – одна из разновидностей популярной операционной системы, которая является инструментом, преимущественно направленным на использование опытными пользователями, т.к. его основная задача – проведение тестов для выявления уязвимостей, а также проведение мер по их устранению. Kali Linux – специализированная версия дистрибутива, которая позволяет пользователям проводить тесты на проникновение. Широко применяется как хакерами, так и теми, кто с ними борется. Благодаря более чем 300 встроенным инструментам пользователи смогут осуществлять тестирование на проникновение.

Для отслеживания сетевой активности на нашей виртуальной машине используется виртуальный монитор – EtherApe (рис. 1). Сетевой монитор EtherApe даёт наглядное графическое представление того, какая часть полосы пропускания задействована, отображая не только соединения, но и "поток" по каждому соединению, вид протокола по номеру порта и сетевую активность различных хостов.

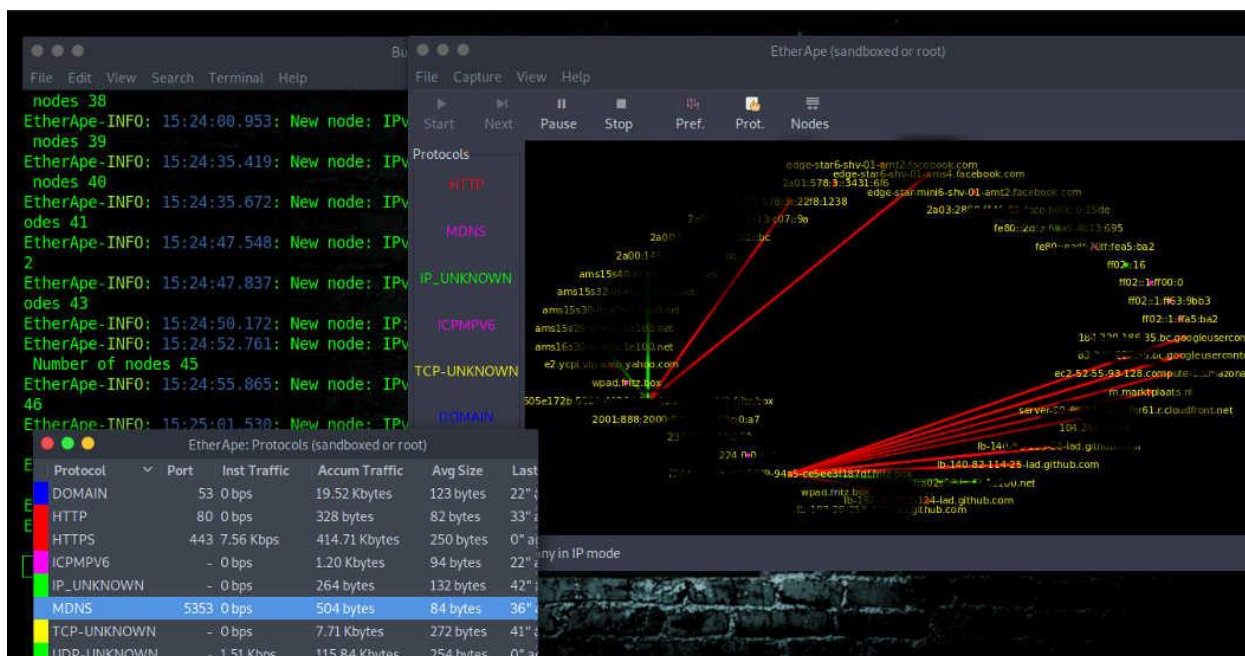


Рис. 1. Сетевой монитор EtherApe.

Для возможности моделирования атак необходимо установить связь с устройством. Данная задача решена с помощью протокола SSH. SSH (англ. Secure Shell — безопасная оболочка) — протокол терминального доступа к *nix-системам, позволяющий производить удалённое управление данными пользователя, запускать различные

служебные команды, работать в консольном режиме с базами данных. С помощью данного протокола, мы сможем имитировать ситуацию подключения к нашему макету со стороннего устройства.

С помощью запуска специальной утилиты Kali-Linux – Hydra. Hydra - это программное обеспечение с открытым исходным кодом для перебора паролей в реальном времени различных онлайн-сервисов, веб-приложений, FTP, SSH и других протоколов. Особенность инструмента в том, что здесь выполняется перебор не по хэшу, а напрямую с помощью запросов к серверу, что позволяет протестировать систему защиты на определение атак подбора пароля.

Демонстрация работы комплекса

С помощью утилиты Hydra была проведена атака грубой силы на макете устройства Интернета Вещей, с помощью команды в терминале (рис. 2а). В результате начинается перебор пары логин + пароль, в нашем случае pi + пароль из словаря. В результате, через несколько минут вывелась пара логина и пароля, соответствующие нашему устройству (рис. 2б).

```
root@kali:~# hydra -V -f -t 4 -l pi -P /home/daniil/Documents/500-worst-passwords.txt ssh://192.168.0.106
```

а)

```
[ATTEMPT] target 192.168.0.106 - login "pi" - pass "bitches" - 268 of 501 [child 0] (0/0)
[ATTEMPT] target 192.168.0.106 - login "pi" - pass "green" - 269 of 501 [child 2] (0/0)
[ATTEMPT] target 192.168.0.106 - login "pi" - pass "super" - 270 of 501 [child 3] (0/0)
[ATTEMPT] target 192.168.0.106 - login "pi" - pass "qazwsx" - 271 of 501 [child 1] (0/0)
[ATTEMPT] target 192.168.0.106 - login "pi" - pass "magic" - 272 of 501 [child 2] (0/0)
[ATTEMPT] target 192.168.0.106 - login "pi" - pass "lakers" - 273 of 501 [child 3] (0/0)
[ATTEMPT] target 192.168.0.106 - login "pi" - pass "rachel" - 274 of 501 [child 1] (0/0)
[ATTEMPT] target 192.168.0.106 - login "pi" - pass "slayer" - 275 of 501 [child 0] (0/0)
[ATTEMPT] target 192.168.0.106 - login "pi" - pass "raspberrry" - 276 of 501 [child 2] (0/0)
[22][ssh] host: 192.168.0.106 login: pi password: raspberrry
[STATUS] attack finished for 192.168.0.106 (valid pair found)
1 of 1 target successfully completed, 1 valid password found
Hydra (https://github.com/vanhauser-thc/thc-hydra) finished at 2020-05-18 18:01:18
root@kali:~#
```

б)

Рис. 2. Выполнение (а) и результат работы (б) утилиты Hydra из состава Kali Linux для brute-force атаки.

Для моделирования атаки типа Shell-injection в макете управляющего Web-интерфейса была реализована соответствующая уязвимость. Пример запроса, вызывающего перезагрузку устройства, приведен Синтаксис запроса представлен на рис. 3 непосредственно в форме для ввода нового имени хоста:

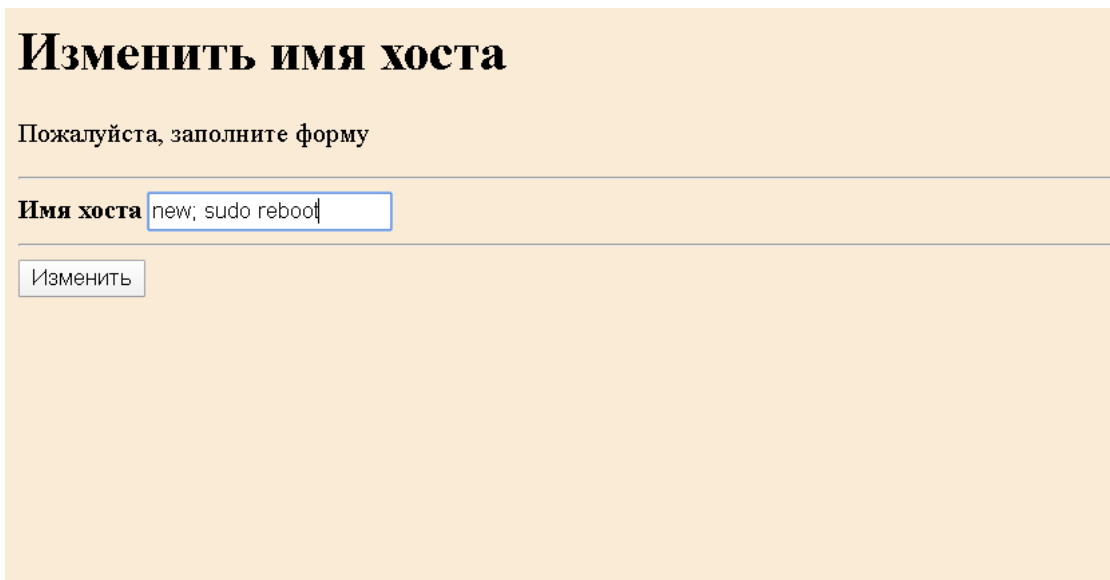


Рис. 3. Проведение shell-injection для перезагрузки устройства.

В данном случае, после выполнения данной команды наше устройство будет перезагружено. Если бы мы использовали команду `–sudo halt`, то наше устройство просто выключилось бы. В результате, веб-сервер нашего устройства перестал функционировать (рис. 4).

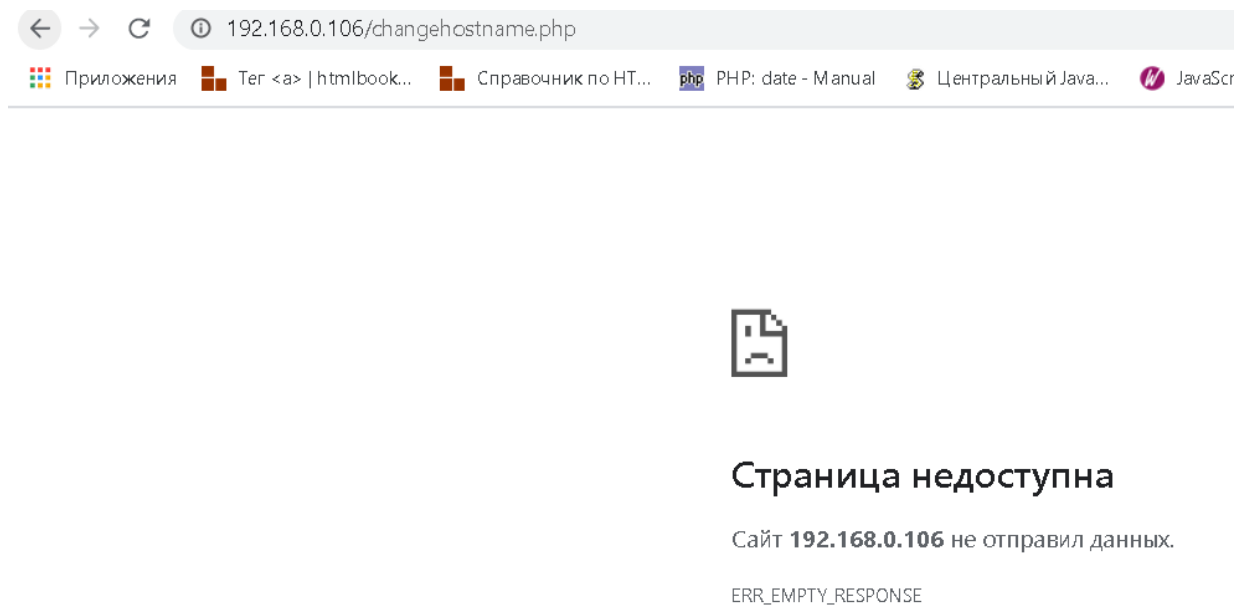


Рис. 4. Результат shell-injection при выключении устройства.

Заключение

Представленная работа демонстрирует возможности разработанного комплекса для моделирования компьютерных атак на модель устройства Интернета вещей. С помощью данного комплекса планируется собирать данные для машинного обучения детектора атак, а также исследовать эффективность работы детектора.

Библиографический список

1. Информационная безопасность интернета вещей [Электронный ресурс]// URL: <https://center2m.ru/informatsionnaya-bezopasnost-veschey>
2. Ботнеты в 2019 году: актуальные опасности и вред, наносимый малому и среднему бизнесу [Электронный ресурс]//URL: <https://habr.com/ru/company/zyxel/blog/449186/>
3. Жилинский Д.И., Елисеев В.Л. Разработка типового устройства Интернета Вещей для изучения компьютерных атак и методов защиты от них // Сборник трудов XXIX Международной научно-технической конференции «Современные технологии в задачах управления, автоматизации и обработки информации», Алушта, Республика Крым, Россия, 14–20 сентября 2020 г., с.161–162.
4. Жилинский Д.И., Елисеев В.Л. Разработка подсистемы подготовки данных для обнаружения атак на устройство интернета вещей (в печати).
5. Степашкин М.В., Котенко И.В., Богданов В.С. Моделирование атак для активного анализа уязвимостей компьютерных сетей // Вторая всероссийская научно-практическая конференция по имитационному моделированию и его применению в науке и промышленности «Имитационное моделирование. Теория и практика» ИММОД-2005. С.269-273.

D.I. Zhilinskiy, V.L. Eliseev

COMPLEX FOR MODELING ATTACKS ON A TYPICAL DEVICE OF THE INTERNET OF THINGS

National Research University “Moscow Power Engineering Institute”

The work is devoted to modern problems of information security of smart devices connected to the Internet. The analysis of the most famous computer epidemics for such devices is carried out. Typical attacks are identified and an approach to their detection is presented. A software complex has been developed for simulating attacks in order to develop and test the protection system.

Keywords: network attacks, IoT devices, modeling complex

ЭМПИРИЧЕСКИЙ КРИТЕРИЙ КАЧЕСТВА ОДНОКЛАССОВОГО КЛАССИФИКАТОРА

¹Национальный исследовательский университет «МЭИ»,

²АО «ИнфоТеКС»

Рассматривается проблема синтеза и оценки качества одноклассовых классификаторов. Проводится анализ общепризнанных метрик качества классификации и отмечается их неудовлетворительность для задачи одноклассовой классификации. Вводится новый критерий качества для одноклассовой классификации, включающий четыре метрики, и демонстрируется его применение на примере нейросетевого автокодировщика.

Ключевые слова: метрики качества, одноклассовая классификация, автокодировщик, искусственные нейронные сети.

Введение

В настоящее время значительные усилия исследователей всего мира направлены на развитие методов искусственного интеллекта, в частности, машинного обучения, что позволяет во многих случаях решать задачи, аналогичные решаемым человеком. Многие из методов машинного обучения посвящены решению задачи классификации: одноклассовой и многоклассовой. В частности, для решения этих задач применяются машины опорных векторов (SVM), метод K-средних, решающие деревья и искусственные нейронные сети. Современные подходы к машинному обучению классификаторов на основе той или иной формальной математической структуры основано на предположении о том, что данные, доступные на этапе обучения, достаточно хорошо соответствуют всем возможным данным, с которыми в дальнейшем предполагается работать.

Контроль за соблюдением справедливости этого предположения для многоклассовых классификаторов основан на анализе матрицы ошибок (confusion matrix) и ряда метрик, рассчитываемых на её основе [1]. В случае одноклассовой классификации эти метрики неприменимы по ряду причин, что затрудняет оценку качества одноклассового классификатора. Это снижает доверие к результатам и делает менее предсказуемым применение одноклассового классификатора в практических задачах. Между тем, одноклассовая классификация представляет собой чрезвычайно ценный инструмент машинного обучения, так как классификатор может настраиваться по неразмеченным данным, что в ряде случаев является критически важным. Рассмотрим задачу оценки качества одноклассового классификатора более подробно.

Обзор литературы

Методология, основанная на убежденности в хорошей представительности обучающей выборки по отношению к генеральной совокупности, в последние десятилетия привела не только к множеству эффективных результатов, таких как распознавание рукописных символов, идентификации людей по фотографиям, но и к ряду неожиданных эффектов некорректной классификации. Во-первых, выяснилось, что за пределами обучающей выборки классификаторы ведут себя непредсказуемо [2]. Во-вторых, оказалось возможным формировать примеры для некорректной классификации целенаправленно [3].

Для многоклассовой классификации существует отработанная методология оценки качества, основанная на анализе матрицы ошибок. В задаче классификации двух классов (binary classification) матрица ошибок имеет вид, показанный в таблице 1, причем в ячейках таблицы указывается количество фактов корректной (True) и некорректной (False) классификации для каждого из классов, традиционно называемых положительным (Positive) и отрицательным (Negative).

Матрица ошибок

Класс Классификация	Positive	Negative
Positive	TP	FP
Negative	FN	TN

Оценка качества классификации делается на основе метрик, рассчитываемых по следующим хорошо известным формулам:

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN}$$

$$Precision = \frac{TP}{TP + FP}$$

$$Recall = \frac{TP}{TP + FN}$$

$$F_{\beta} = (1 + \beta^2) \frac{Precision \cdot Recall}{(\beta^2 \cdot Precision) + Recall}$$

При $\beta = 1$ F-мера превращается в среднее гармоническое:

$$F_1 = 2 \frac{Precision \cdot Recall}{Precision + Recall}$$

Кроме этих метрик некоторыми исследователями вводятся и другие, позволяющие более эффективно характеризовать те или иные аспекты алгоритмов машинного обучения, например, [1] и [4]. Также для оценки качества используется так называемые ROC-кривые [5]. Следует отметить, что все метрики основаны на классах, описываемых ограниченными счётными множествами, что хорошо согласуется с практическими реализациями, оперирующими, тестовыми и контрольными эмпирическими выборками. Другой особенностью большинства метрик является их ориентация на число рассматриваемых классов не менее двух.

С точки зрения традиционных подходов задача одноклассовой классификации выглядит как предельно несбалансированная бинарная классификация [6]. Ранее отмечалось, что проблема несбалансированности может быть устранена при условии задания специфической структуры классификатора [7]. В обзоре [8] отмечается, что имеет место сложность с объективной оценкой ошибки классификации, а также неоднозначность понятия «границы класса». Фактически для одноклассовой классификации возможно использовать только те метрики бинарной классификации, которые основаны на эмпирически доступных показателях **TP** и **FN**, например, **Recall**. Для вычисления **Precision** используется адаптированное для единственного класса определение [9]:

$$Precision = \frac{TP}{TP + FN}$$

Мы видим, что в формуле искусственно заменен показатель **FP** на **FN**. С одной стороны, это позволило определить точность классификации как отношение корректно классифицированных примеров к общему числу, с другой – теперь точность (**Precision**) уже не может использоваться для оценки влияния некорректных фактов позитивной классификации **FP**. Между тем, некорректная позитивная классификация является значительной проблемой всех алгоритмов машинного обучения. Для одноклассовых

классификаторов, используемых для обнаружения новизны, некорректное отнесение большого числа примеров к норме делает бессмысленным применение детектора аномалий.

Можно заключить, что проблема оценки качества одноклассовых классификаторов известна исследователям, однако эффективные специфические метрики для таких классификаторов не были введены.

Методология

Основным назначением какой-либо метрики качества является оценка соответствия некоторому идеалу. Рассмотрим одноклассовый классификатор, как способ верного установления факта принадлежности или непринадлежности объекта к заданному множеству. Заметим, что несмотря на работу с ограниченными счётными множествами на этапе обучения, применение классификаторов предполагается для любых объектов, чьё описание позволяет формально проводить классификацию. Таких объектов может быть не только значительно больше, чем имелось на этапе обучения: их может быть несчётное количество, например, если целевое множество определено непрерывно. Определим одноклассовый классификатор на непрерывном множестве. Для этого дадим несколько определений.

Обозначим через X множество объектов, для которых будем рассматривать задачу классификации. Для одноклассового классификатора выходом будет $Y = \{0,1\}$, причем 1 обозначает принадлежность к целевому классу, а 0 – непринадлежность. Подмножество, соответствующее целевому классу, обозначим как $X_T \subset X$. Зададим функцию принадлежности объекта целевому классу:

$$\forall x \in X: f(x) = \begin{cases} 1, & x \in X_T \\ 0, & x \notin X_T \end{cases}$$

Данная функция является идеальным одноклассовым классификатором, однако на практике, как правило, нет возможности получить все значения $x \in X_T$ и использовать их для обучения. Обычно доступно ограниченное счётное обучающее множество $\hat{X}_T \subset X_T$, из которого формируется обучающая выборка, а также выборки для контроля процесса обучения и тестирования. Обозначим обученный одноклассовый классификатор $CL^1(x): X \rightarrow Y$.

Несмотря на то, что CL^1 обучается на $\hat{X}_T \subset X_T$, в итоге синтеза возможны разнообразные ситуации, соответствующие исходам бинарной классификации (табл. 2):

Таблица 2

Исходы бинарной классификации		
Класс \ Классификация	Positive $x \in X_T$	Negative $x \notin X_T$
Positive	$TP: CL^1(x) = 1$	$FP: CL^1(x) = 1$
Negative	$FN: CL^1(x) = 0$	$TN: CL^1(x) = 0$

Фактически реальный одноклассовый классификатор отличается от идеального тем, что множество объектов, которые он относит к целевому классу, деформировано, то есть, отличается от целевого $X_D \neq X_T$:

$$\forall x \in X: CL^1(x) = \begin{cases} 1, & x \in X_D \\ 0, & x \notin X_D \end{cases}$$

Наглядно подобную ситуацию можно изобразить на рисунке:

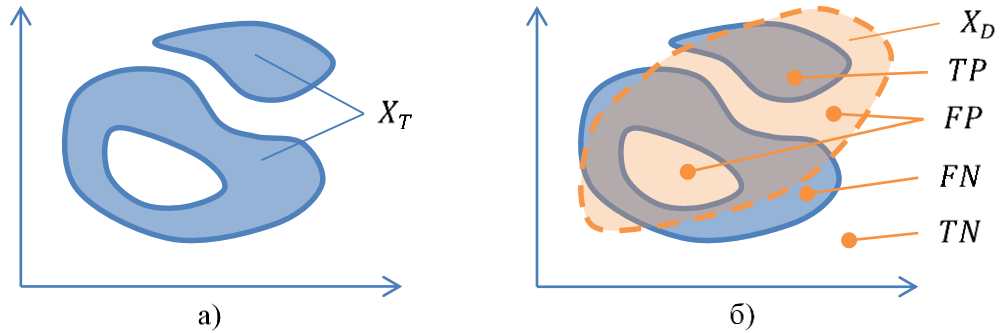


Рис. 1. Область целевого класса (а) и деформированная область (б)

Качеством одноклассового классификатора является близость CL^1 к идеальному f , которая может быть выражена через свойства множества X_D по сравнению с множеством X_T . В частности, идеальности мешают исходы классификации FP и FN . Ситуации FP соответствует случай $X_D \setminus X_T \neq \emptyset$, а ситуации FN – случай $X_T \setminus X_D \neq \emptyset$. Для непрерывных множеств объектов можно ввести понятие объёма, для чего воспользуемся обозначением $|X|$, причём $|\emptyset| = 0$. В этом случае можно описать наличие ситуаций FP и FN через объём так:

$$FP: |X_D \setminus X_T| \geq 0$$

$$FN: 0 \leq |X_T \setminus X_D| \leq |X_T|$$

Исход корректной классификации описывается через объём:

$$TP: 0 \leq |X_T \cap X_D| \leq |X_T|,$$

причём из $|X_T \cap X_D| = |X_T|$ следует $|X_D \setminus X_T| = 0$ и $|X_T \setminus X_D| = 0$.

Желательно, чтобы метрика качества классификатора не зависела от объёма целевого множества, что позволит сравнивать качество классификаторов, обученных для различных прикладных задач. Введем нормированные на объём целевого множества метрики:

$$Excess = \frac{|X_D \setminus X_T|}{|X_T|} \quad (1)$$

$$Deficit = \frac{|X_T \setminus X_D|}{|X_T|} \quad (2)$$

$$Coating = \frac{|X_T \cap X_D|}{|X_T|} \quad (3)$$

Метрика *Excess* (англ. избыток) принимает значение больше 0 в том случае, если CL^1 неправильно классифицирует объекты за пределами целевого множества. Метрика *Deficit* (англ. недостаток) принимает значение от 0 до 1, причём если CL^1 неправильно классифицирует объекты в пределах целевого множества, то метрика больше 0. Предельным значением $Deficit = 1$ является ситуация, когда целевое и классифицированное множества не пересекаются. Метрика *Coating* принимает значения от 0 до 1, причём значению 0 соответствует нулевое перекрытие множеств, а 1 соответствует полное покрытие целевого множества.

Точность аппроксимации целевого множества может быть оценена как отношение площади целевого множества к площади деформированного классификатором:

$$Approx = \frac{|X_T^*|}{|X_D^*|} \quad (4)$$

Идеальный одноклассовый классификатор характеризуется следующими значениями метрик:

$$Excess = 0, Deficit = 0, Coating = 1, Approx = 1$$

Рассмотрим способ вычисления введенных метрик для задачи оценки качества одноклассового классификатора для дискретных множеств, с которыми приходится иметь

дело в практических задачах машинного обучения. Как уже отмечалось, традиционно используются метрики, основанные на количестве примеров в выборках. В противоположность этому, введем понятие дискретного объема выборки. Для этого введем в векторном пространстве R^n , в котором обычно представляются объекты для классификации, разбиение на атомарные элементы. Это может быть равномерная по каждой из осей координат сетка или иной способ разбиения. Таким образом, любой точке x в этом пространстве можно соотнести ячейку дискретного разбиения a :

$$\forall x \in R^n \exists! a | x \in a$$

Будем аппроксимировать множества, введенные ранее, на дискретизированное пространство. Таким образом, перейдем от непрерывных множеств X_T и X_D к дискретным множествам X_T^* и X_D^* . Для этих множеств объем будет равен сумме объемов элементов разбиения, образующих эти множества. В случае разбиения на ячейки одинакового объема, на единичный объем элемента разбиения.

Поскольку на практике нам недоступно полное множество X_T , вместо которого имеется только конечное счётное обучающее подмножество $\hat{X}_T \subset X_T$, то дискретное множество X_T^* образуется из элементов разбиения, содержащих хотя бы одну точку из \hat{X}_T :

$$X_T^* = \{a | \exists x \in \hat{X}_T \& x \in a\}$$

Дискретное множество X_D^* образуется из элементов разбиения, какая-либо точка внутри которых дает положительный результат классификации:

$$X_D^* = \{a | \exists x \in a : CL^1(x) = 1\}$$

Для расчета метрик *Excess*, *Deficit* и *Coating* для дискретизированных множеств необходимо определить процедуру расчета их объемов. Объем $|X_T^*|$ рассчитывается через сумму объемов атомарных элементов разбиения, образующих множество X_T^* . Объем $|X_D^* \setminus X_T^*|$ определяется как сумма объемов ячеек, входящих только в X_D^* , но не входящих в X_T^* . Объем $|X_T^* \setminus X_D^*|$ определяется как сумма объемов ячеек, входящих только в X_T^* , но не входящих в X_D^* . Объем $|X_T^* \cap X_D^*|$ рассчитывается как сумма объемов ячеек, входящих как в X_T^* , так и в X_D^* .

Эксперименты

Для наглядной демонстрации введенных метрик качества одноклассовой классификации были проведены эксперименты с обучением нейросетевого одноклассового классификатора – автокодировщика (autoencoder). Для этого было взято множество \hat{X}_T случайно сгенерированных точек в двумерной области с границами, образующими кольцо, синтезированы автокодировщики АЕ1 и АЕ2 с разной нейросетевой архитектурой и проведены расчеты введенных метрик качества *Excess*, *Deficit* и *Coating*.

Первый автокодировщик АЕ1 простой архитектуры с 2 нейронами во входном и выходном слое и тремя скрытыми слоями с 3, 1 и 3 нейронами соответственно был обучен с помощью алгоритма оптимизации ADAM на протяжении 50000 эпох. В результате классификации множества точек просканированного пространства был выявлен ограниченный набор точек, отнесенных обученным автокодировщиком к известному классу и образующих деформированное множество \hat{X}_D . Обучающий набор \hat{X}_T , а также множество точек, распознанных автокодировщиком, \hat{X}_D изображены на рисунке 2. Дискретные области X_T^* и X_D^* , включающие обучающее \hat{X}_T и деформированное множество \hat{X}_D , были получены благодаря разбиению исследуемой части пространства на атомарные объекты – прямоугольники с известной длиной и шириной – 0.27. Каждый атомарный объект является частью дискретной области только в случае наличия в нём хотя бы одной точки известного множества. Дискретные области X_T^* и X_D^* выделены на рисунке 2 цветом.

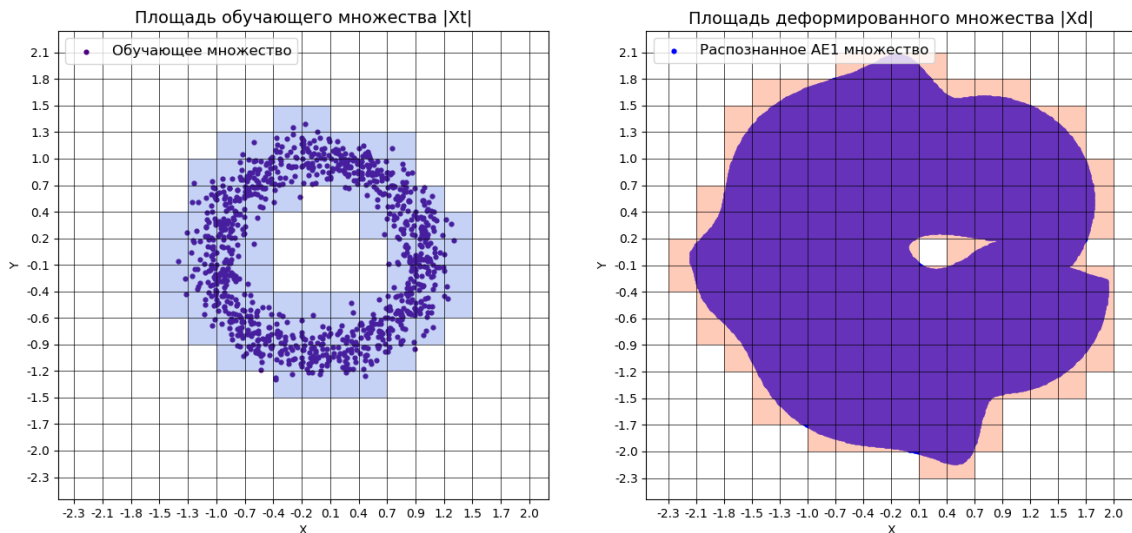


Рис. 2. Дискретные области обучающего X_T^* и деформированного множества X_D^* (AE1)

Таким образом, появляется возможность рассчитать объем или в двумерном случае - площадь дискретной области, и оценить насколько деформированная область не соответствует идеалу – области обучающего множества. Площадь одного элемента разбиения равна произведению длины и ширины: 0.073, а сумма элементов разбиения, составляющих интересующую область, определяет площадь. Например, дискретная область обучающего множества X_T^* состоит из 78 элементов разбиения, а деформированная X_D^* - из 188. Таким образом, для автокодировщика AE1 результат соотношения площадей составляет $Approx = 0.41$, что уже говорит о неидеальности классификатора, реализуемого этим автокодировщиком.

Для расчета метрики *Excess* необходимо выявить дискретную область деформированного множества X_D^* , не включающую дискретную область обучающего множества X_T^* , и рассчитать площадь: $|X_D^* \setminus X_T^*|$. Область $X_D^* \setminus X_T^*$ выделена на первом графике рисунка 3 штриховкой. Таким образом, величина $Excess=1.4$. Поскольку для идеального классификатора *Excess* должен быть равен 0, то можно утверждать, что согласно данной метрике автокодировщик будет давать существенное количество ложноположительных срабатываний (*FP*).

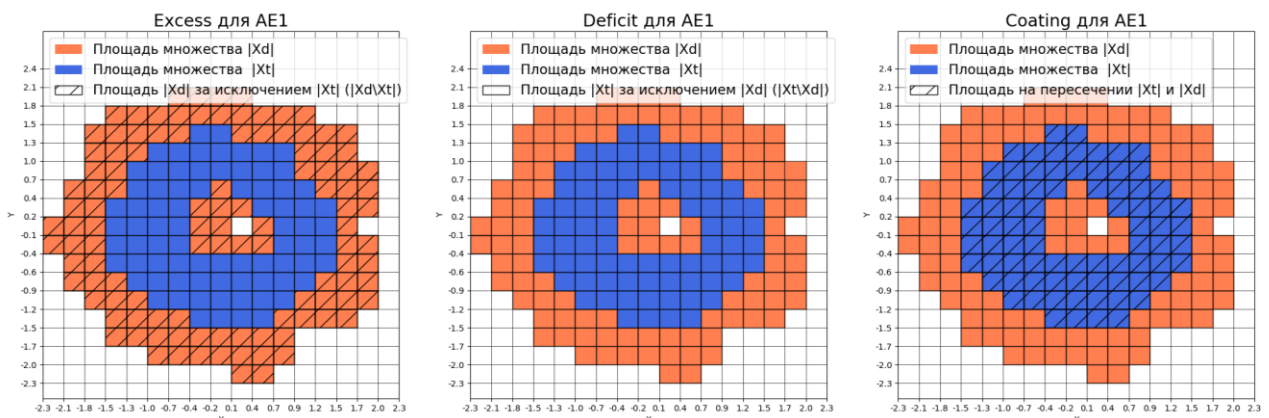


Рис. 3. Визуализация дискретных областей, необходимых для расчета метрик *Excess*, *Deficit* и *Coating* (AE1)

Рассмотрим второй график на рисунке 3, необходимый для понимания расчета метрики *Deficit*, заключающегося согласно формуле (2) в нахождении отношения площади дискретной области X_T^* , не включающей область X_D^* : $X_T^* \setminus X_D^*$, к площади области обучающего множества X_T^* . По графику становится понятно, что такой области нет и соответственно её площадь равна 0. В таком случае метрика *Deficit* = 0 для автокодировщика АЕ1 имеет идеальное значение.

На третьем графике рисунка 3 штриховкой обозначена дискретная область, соответствующая пересечению областей обучающего множества и деформированного множества $X_T^* \cap X_D^*$. Как видно, она целиком совпадает с областью обучающего множества X_T^* и отношение этих площадей, согласно формуле (3) определяет величину метрики *Coating* = 1, что также соответствует идеальному случаю.

Таким образом, для автокодировщика АЕ1 метрики будут равны:

$$Excess = 1.4, Deficit = 0, Coating = 1, Approx = 0.41$$

Достоверно известно, что в ряде случаев усложнение архитектуры нейронной сети (увеличение количества скрытых слоев и/или увеличение числа нейронов в них) позволяет добиться более высокого качества классификации. Усложним архитектуру нейросети для того, чтобы улучшить качество классификации. В архитектуре автокодировщика АЕ2 число нейронов в двух крайних скрытых слоях равно 5. Алгоритм и длительность обучения автокодировщика АЕ2 полностью совпадают с установленными для автокодировщика АЕ1. Следует предположить, что качество классификации и метрики должны стать ближе к идеальному одноклассовому классификатору.

Результат классификации – деформированное множество X_D^* для автокодировщика АЕ2 – представлен на рисунке 4.

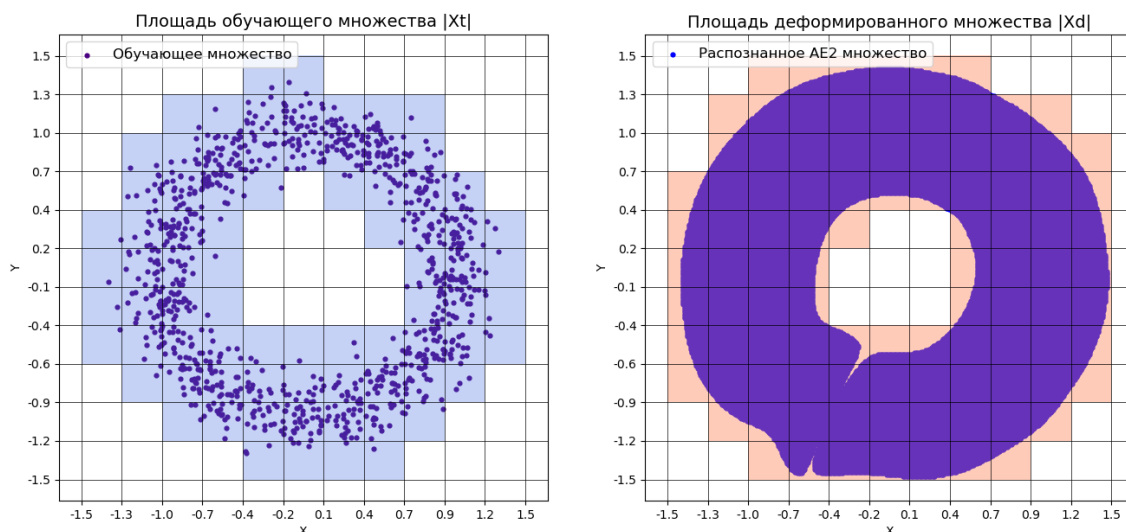


Рис. 4. Дискретные области обучающего X_T^* и деформированного множества X_D^* (АЕ2)

Рассчитаем значения метрик для автокодировщика АЕ2, воспользовавшись формулами 1-4 и визуализацией областей (рис. 5):

$$Excess = 0.27, Deficit = 0, Coating = 1, Approx = 0.78$$

Метрики подтверждают предположение о более высоком качестве классификатора, построенного на основе автокодировщика АЕ2 более сложной архитектуры, чем АЕ1.

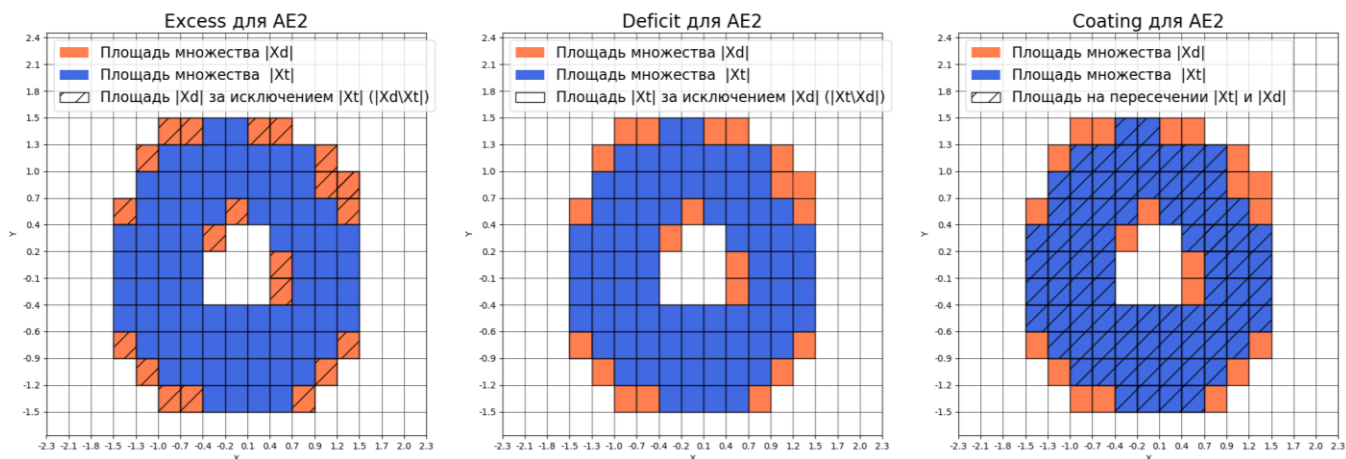


Рис. 5. Визуализация дискретных областей, необходимых для расчета метрик Excess, Deficit и Coating (AE2)

Результаты проведенных экспериментов в виде расчетных значений предложенных метрик оценки качества одноклассовой классификации автокодировщиком простой архитектуры AE1 и автокодировщиком усложненной архитектуры AE2 сведены в таблицу 3. Также, для сравнения в таблице приведены предельные значения метрик, соответствующие идеальному одноклассовому классификатору.

Таблица 3

Значения метрик Excess, Deficit и Coating для AE1 и AE2

	<i>AE1</i>	<i>AE2</i>	<i>Идеальный CL^1</i>
<i>Approx</i>	0.41	0.78	1
<i>Excess</i>	1.4	0.27	0
<i>Deficit</i>	0	0	0
<i>Coating</i>	1	1	1

Анализ полученных значений подтверждает полезность введенных метрик для оценки качества одноклассовых классификаторов, а также соответствует субъективному пониманию качества классификации.

Обсуждение

Предложенные метрики качества разработаны специально для одноклассовых классификаторов, с учетом и целью устранить недостатки, свойственные общеизвестным критериям качества многоклассовой классификации. Предложенный подход уникален и ценен тем, что оперирует не количеством примеров в выборке, а областями, покрываемыми точками обучающего множества. Введенные метрики позволяют оценивать качество классификации в многомерных пространствах, когда невозможно визуализировать области обучающих и классифицируемых точек.

Метрики применимы к одноклассовым классификаторам любых видов: нейросетям, машинам опорных векторов, решающим деревьям и т.д. Можно предположить, что использование метрик позволит повысить достоверность результатов и доверие к одноклассовой классификации. Данные метрики также можно использовать для сравнения и целенаправленной оптимизации классификаторов.

К объективным недостаткам подхода можно отнести следующие особенности:

- экспоненциальный рост вычислительной сложности и потребляемой памяти с ростом размерности пространства X ;

- высокая вычислительная сложность сканирования пространства для определения деформированного множества X_D^* , растущая экспоненциально с уменьшением шага сетки разбиения;
- экспоненциально растущий с уменьшением шага сетки объём памяти для выполнения операций над дискретизированными множествами;
- зависимость величин рассчитываемых метрик от шага сетки.

В то же время, можно отметить, что вычислительная сложность определения X_T^* линейно зависит от размерности обучающего множества \hat{X}_T .

Заключение

Была рассмотрена проблема оценки качества одноклассовой классификации и сформулирован критерий на основе теоретико-множественных операций в предположении о непрерывности целевого и классифицированного множеств. Предложены новые метрики качества Approx, Excess, Deficit, Coating, основанные на оценке объемов множеств. Продемонстрировано применение введенных метрик для двух нейросетевых автокодировщиков разной архитектуры. Подтверждена хорошая интерпретируемость значений метрик, а также субъективное соответствие метрик ожидаемым результатам сравнения двух автокодировщиков.

Библиографический список

1. Forman G., "An extensive empirical study of feature selection metrics for text classification," *Journal of Machine Learning Research (JMLR)*, no. 3, 2003.
2. Haley P. and Soloway D., "Extrapolation limitations of multilayer feedforward neural networks," in *IJCNN International Joint Conference on Neural Networks*, Baltimore, MD, USA, 1992.
3. Szegedy C. et al., "Going deeper with convolutions," in *2015 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, Boston, MA, 2015.
4. Powers D., "Evaluation: From Precision, Recall and F-Factor to ROC, Informedness, Markedness & Correlation," *Technical Report SIE-07-001*, December 2007.
5. Fawcett T., "An Introduction to ROC Analysis," *Pattern Recognition Letters*, vol. 27, no. 8, p. 861–874, 2006.
6. Nguyen G., A. Bouzerdoum and S. Phung, "Learning Pattern Classification Tasks with Imbalanced Data Sets," *Pattern Recognition*, no. 10, 2009.
7. Lee W. and Liu B., "Learning with positive and unlabeled examples using weighted logistic regression," in *In Proceedings of the Twentieth International Conference on International Conference on Machine Learning (ICML'03)*, 2003.
8. Khan S. and Madden M., "One-class classification: Taxonomy of study and review of techniques," *The Knowledge Engineering Review*, vol. 29, no. 3, pp. 345-374, 2014.
9. Бурмистров М. and Сандуляну Л., "Вероятностная модель одноклассовой классификации," *Машинное обучение и анализ данных*, vol. 1, no. 4, pp. 420-427, 2012.

EMPIRICAL CRITERION FOR THE QUALITY OF A ONE-CLASS CLASSIFIER

¹ *National Research University "Moscow Power Engineering Institute",*

² *JSC Infotecs*

The problem of synthesis and quality assessment of one-class classifiers is considered. The analysis of generally accepted metrics of classification quality is carried out and their unsatisfactoriness for the problem of one-class classification is noted. A new quality criterion for one-class classification is introduced, which includes four metrics, and its application is demonstrated using the example of a neural network autoencoder.

Keywords: quality metrics, one-class classification, auto-encoder, artificial neural networks.

ИССЛЕДОВАНИЕ НЕЧЕТКОЙ СИСТЕМЫ УПРАВЛЕНИЯ НЕСТАЦИОНАРНЫМ ОБЪЕКТОМ

¹Национальный исследовательский университет «МЭИ»,

²АО «ИнфоТеКС»

Рассмотрена задача синтеза системы управления с обратной связью для объекта, свойства которого ступенчато изменяются в процессе работы. Исследована система управления с регуляторами двух типов – классическим ПД и нечетким логическим ПД. Представлены результаты имитационного моделирования в пакете SimInTech системы управления с различными сигналами уставки и критериями качества управления. Проведен анализ и выработаны рекомендации по более эффективному применению нечетких логических регуляторов для управления нестационарными объектами.

Ключевые слова: система управления, нечеткий логический регулятор, нестационарный объект, ПД-регулятор.

Введение

В наше время уровень развития промышленности требует при разработке систем автоматического управления техническими объектами комплексного подхода, так как для современного уровня развития технических систем характерно усложнение структуры объектов управления. Это продиктовано, во-первых, необходимостью повышения качества управления при минимальных расходах на создание и эксплуатацию систем, во-вторых, усложнением структуры объекта управления или, например, его нестационарностью, в-третьих, усложнением функций, которые выполняет объект, и, как следствие, увеличением факторов неопределенности, которые необходимо учитывать при управлении объектом.

Большинство технологических процессов и установок в промышленности являются нестационарными объектами управления, то есть их параметры могут изменять во время процесса. [1] Требования к качеству системы управления также постоянно возрастают. В связи с этим появляется необходимость в развитии таких технологий, которые бы позволили обеспечить требуемые показатели качества систем с неполным или неточным математическим описанием.

Существует множество методов классической теории автоматического управления, которые позволяют добиться определённых показателей качества по каким-то критериям, но математические модели, используемые в таких методах, не всегда способны полноценно описать современные системы управления. Именно противоречие между сложностью системы и возможностью её точного описания можно считать основной предпосылкой появления такого математического аппарата, известного как «нечёткая логика». В современных системах управления широкое применение получили нечеткие логические регуляторы, использующие алгоритмы нечеткого логического вывода. Многочисленные варианты построения нечётких систем различаются не только особенностями реализации этапов нечёткого логического вывода, но и схемными решениями. Так, в [6] было предложено различать нечёткие логические регуляторы 1-го и 2-го рода (рис. 1.). В ПИ-, ПД- и ПИД-регуляторах 1-го рода используются параллельно включенные одномерные нечёткие вычислители, сигналы с выходов которых складываются. В регуляторах 2-го рода используется один многомерный нечёткий вычислитель [2].».

Здесь $e(t)$ и $x(t)$ – входной и выходной сигналы НЛР соответственно, s – оператор Лапласа. В ПИ-, ПД- и ПИД-регуляторах 1-го рода используются параллельно включенные одномерные нечёткие вычислители, сигналы с выходов которых

складываются. В регуляторах 2-го рода используется один многомерный нечёткий вычислитель [2].

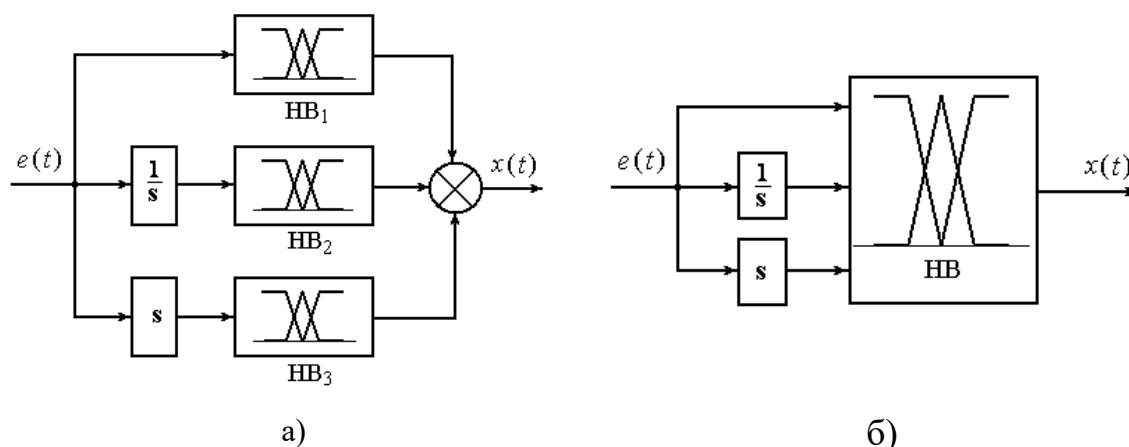


Рис. 1. Нечёткий ПИД-регулятор 1-го рода (а) и 2-го рода (б).

Эти регуляторы могут управлять объектами всех типов, для которых может быть составлена база правил, состоящая из продукций вида «если ..., то...», основанная на знаниях и/или опыте экспертов. Знания экспертов могут быть формализованы с помощью функций принадлежности термов лингвистических переменных и установлением нечеткого соответствия между пространством предпосылок, и пространством заключений. Также для настройки нечеткого логического регулятора является необходимым выбор алгоритма нечеткого логического вывода.

В [7] рассматривается одна из важных задач промышленности: обеспечение качества выпускаемых на предприятиях подшипников. В токарной обработке требуется повышать производительность, сохраняя заданные параметры качества колец подшипника. Одним из факторов, позволяющих увеличить производительность производства, является возможность гибкого управления режимами точения, включая изменение скорости оборотов шпинделя. Однако при этом необходимо обеспечить требуемые характеристики качества управления. Исследования позволили идентифицировать динамические модели станков. В частности, была разработана динамическая модель системы токарного станка, представляющая собой нестационарный объект, меняющий параметры в ходе технологического процесса и существенно зависящая от скорости вращения шпинделя.

Представляется актуальным провести исследование влияния нестационарности объекта на качество управления. Проведем сравнение системы управления нестационарным объектом с помощью нечеткого логического регулятора 1-го рода и с аналогичной системой на основе классического ПД-регулятора. При рассмотрим систему под воздействием уставок различного вида.

Подобные исследования целесообразно проводить с помощью пакетов имитационного моделирования, таких как MATLAB/Simulink и SimInTech. В настоящей работе будет использоваться пакет SimInTech [8], поскольку он обладает всеми необходимыми свойствами, включая библиотеку моделирования блоков нечеткой логики, этот пакет бесплатен для академического использования, имеет русскоязычный интерфейс и обширную справочную систему.

Основная часть

Объект управления

Рассматривается система управления нестационарным динамическим объектом, параметры которого изменяются в ходе технологического процесса.

Допустим, что заранее известны:

- Продолжительность технологического процесса: 100 с
- Время изменения параметров объекта управления: 50 с
- Первоначальные параметры объекта управления (передаточная функция):

$$W_1(s) = \frac{k}{s \cdot (T \cdot s + 1)},$$

где $k = 1, T = 1$.

- Параметры изменившегося объекта (передаточная функция):

$$W_2(s) = \frac{k}{T^2 \cdot s^2 + 2 \cdot \beta \cdot T \cdot s + 1},$$

где $k = 100, T = 0.5, \beta = 0.1$.

В качестве практического примера можно привести изменение частоты оборотов асинхронного двигателя токарного станка. Например, при изменении режима резания на станке меняются параметры объекта управления, а соответственно и частота вращения двигателя станка. Также реверсивные электродвигатели для сверлильных станков являются асинхронными, диапазон их мощности от 1,5 до 15 кВт. Сегодня к способности незначительно изменять частоту вращения при различных внешних воздействиях предъявляются серьезные требования, так же, как и к скорости разгона. Асинхронный двигатель, например, при перегрузке в 2 раза меняет частоту вращения на 10-12%, но этот показатель требуется улучшать. Поэтому современные электродвигатели комплектуют датчиками скорости и сложными системами управления.

Система автоматического управления с нечетким ПД-регулятором 1-го рода

Схема рассматриваемой системы управления с нечетким логическим регулятором первого рода представлена на рис. 2. Она была реализована в пакете Sim In Tech. С помощью блоков библиотеки Нечеткая логика был синтезирован нечеткий регулятор 1-го рода.

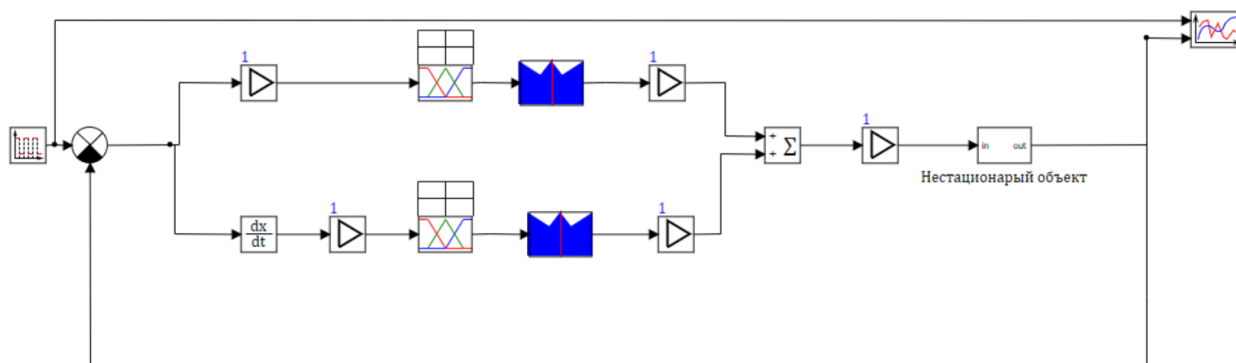


Рис. 2. Нечеткая система управления нестационарным объектом.

Регулятор состоит из двух каналов управления – пропорционального и дифференцирующего с нечёткими вычислителями. До и после каждого нечёткого вычислителя включаются пропорциональные звенья. Сигналы на выходах каналов управления складываются, как и в традиционном линейном ПД-регуляторе.

Для того, чтобы описать работу нечётких вычислителей, были сформулированы лингвистические переменные с терм-множествами {"N", "P", "Z"}, где N – «отрицательное», P – «близкое к нулю», Z – «положительное». Функции принадлежности треугольной формы. База правил состоит из простых правил вида

Π_k : ЕСЛИ «Условие k », ТО «Заключение k ».

База правил (П):

П1: ЕСЛИ $E = "N"$, ТО $X_e = "N"$

П2: ЕСЛИ $E = "Z"$, ТО $X_e = "Z"$

П3: ЕСЛИ $E = "P"$, ТО $X_e = "P"$

П4: ЕСЛИ $D = "N"$, ТО $X_d = "N"$

П5: ЕСЛИ $D = "Z"$, ТО $X_d = "Z"$

П6: ЕСЛИ $D = "P"$, ТО $X_d = "P"$

Существуют различные способы дефаззификации, но в большинстве практических приложений используется метод «центра тяжести»:

$$x^* = \frac{\int_{x_{\min}}^{x_{\max}} x \cdot \mu'(x) dx}{\int_{x_{\min}}^{x_{\max}} \mu'(x) dx}$$

Система автоматического управления с классическим ПД-регулятором

Схема рассматриваемой системы управления с классическим ПД-регулятором представлена на рис. 3 и реализована в пакете SimInTech. С помощью стандартных блоков был синтезирован классический ПД-регулятор.

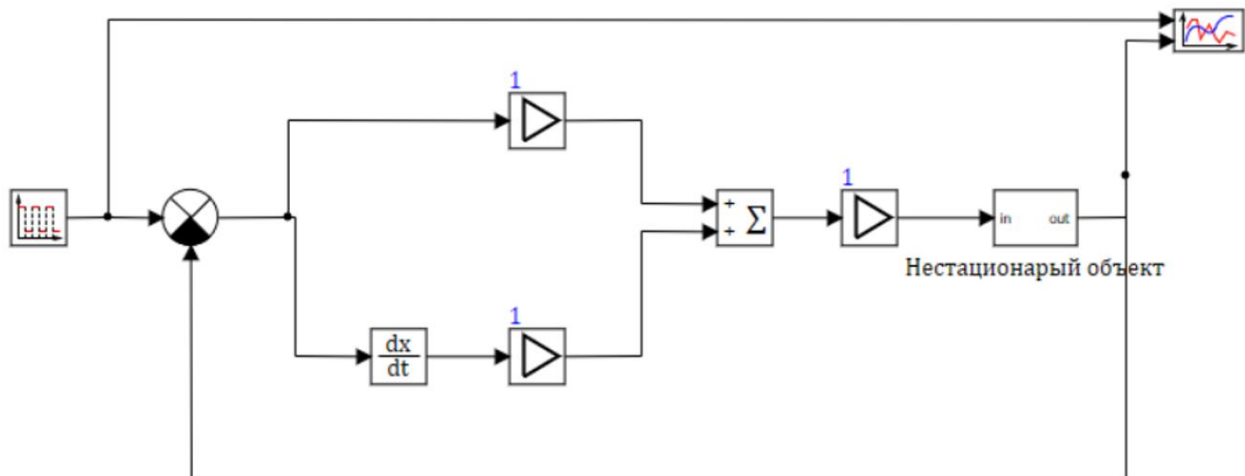


Рис. 3. Схема с ПД-регулятором

Эксперименты

Целью проводимых исследований можно выделить анализ качества управления в условиях нестационарного объекта и последующее сравнение качества управления систем, построенных на традиционном ПД-регуляторе и нечётком ПД-регуляторе 1-го рода. В качестве показателей качества системы рассматривается время регулирования, перерегулирование и среднеквадратичная ошибка.

Было проведено три серии экспериментов с различными входными сигналами:

1. Входной сигнал типа «меандр»
2. Входной сигнал типа «равномерный шум в диапазоне [-1;1]»
3. Входной сигнал типа «нормальный шум»

Во всех экспериментах в момент времени 50 с объект управления изменял свои характеристики с $W_1(s)$ на $W_2(s)$, чем моделировалось изменение скорости вращения шпинделя.

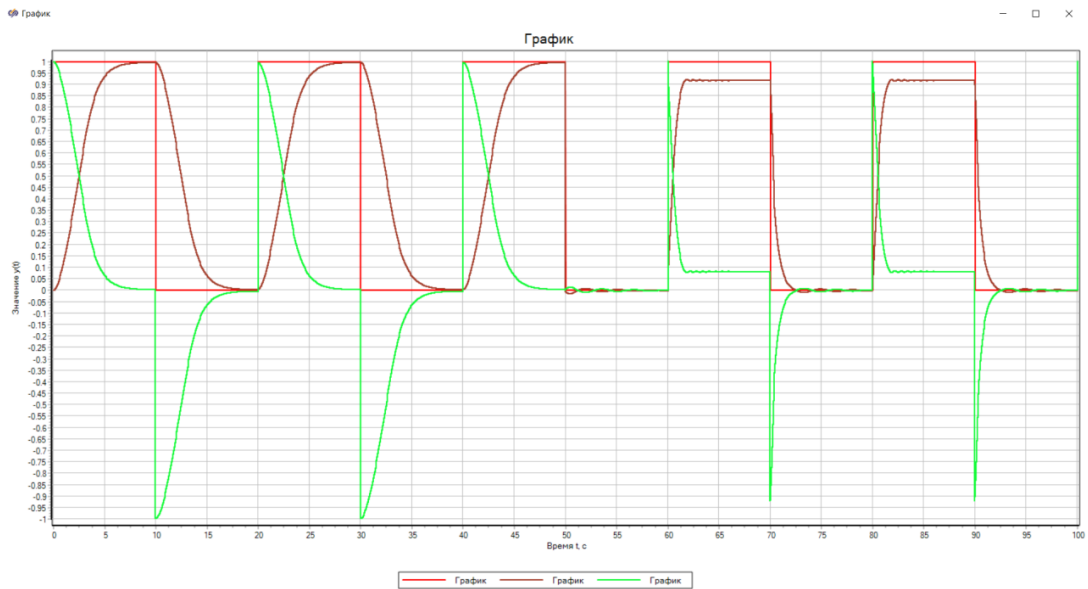


Рис. 4. Переходной процесс нечеткой системы на «меандре»

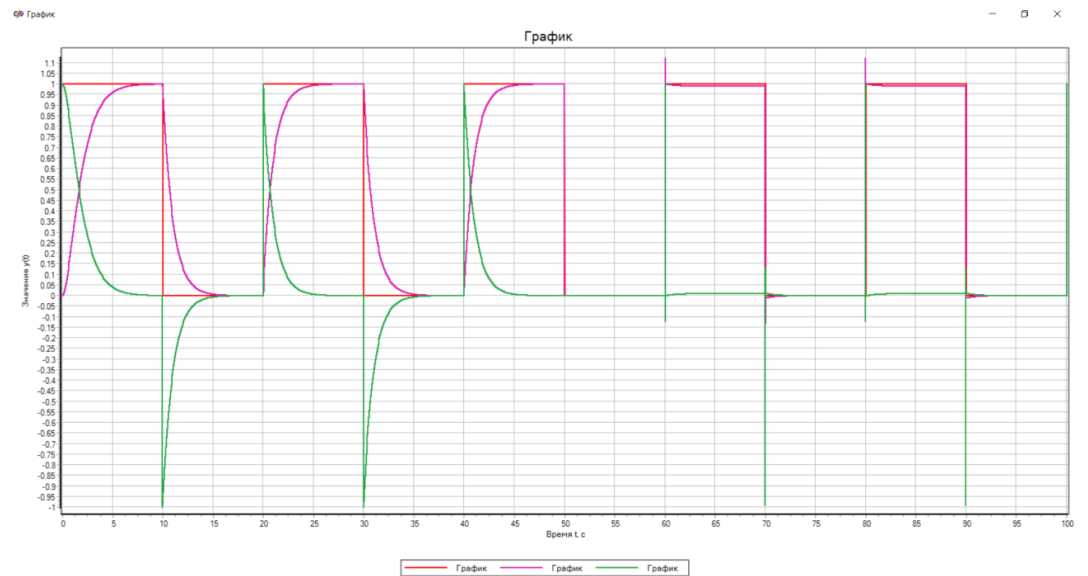


Рис. 5. Переходной процесс системы с ПД-регулятором на «меандре»

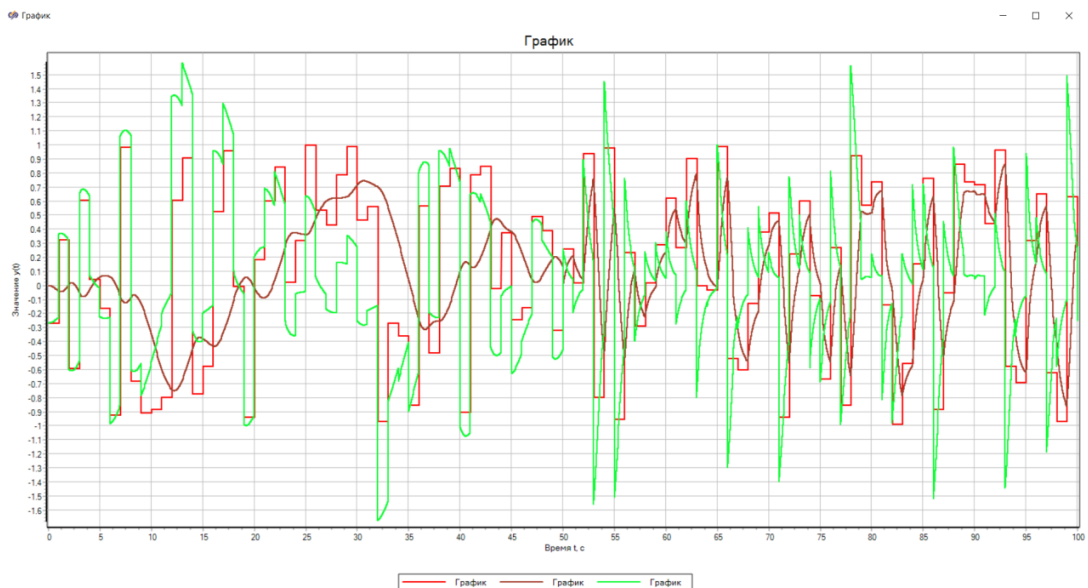


Рис. 6. Переходной процесс нечеткой системы при равномерном шуме

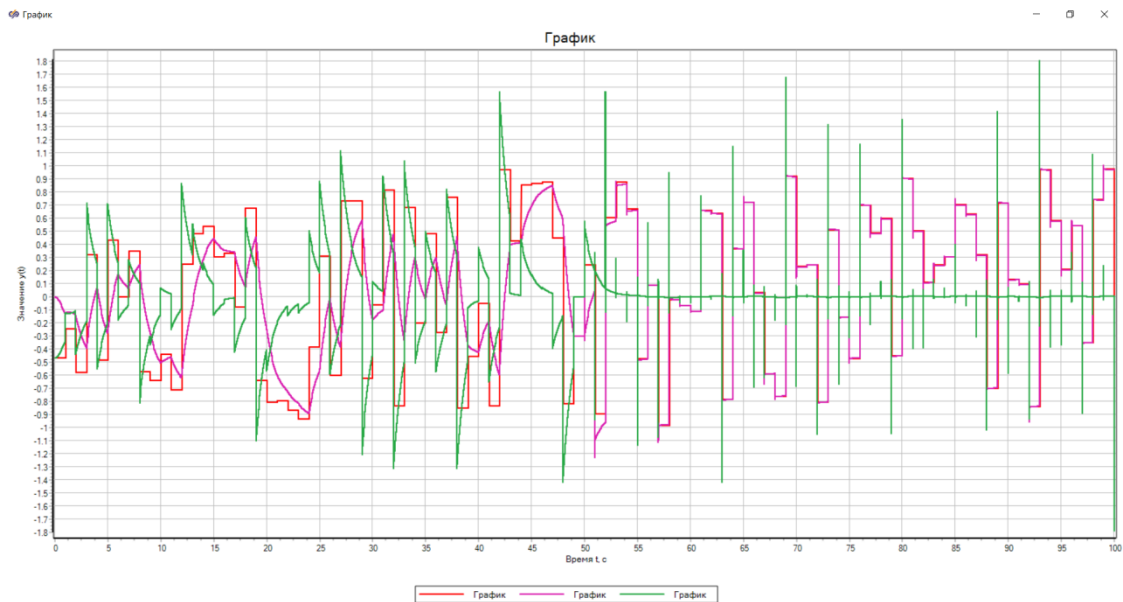


Рис. 7. Переходной процесс классической системы при равномерном шуме

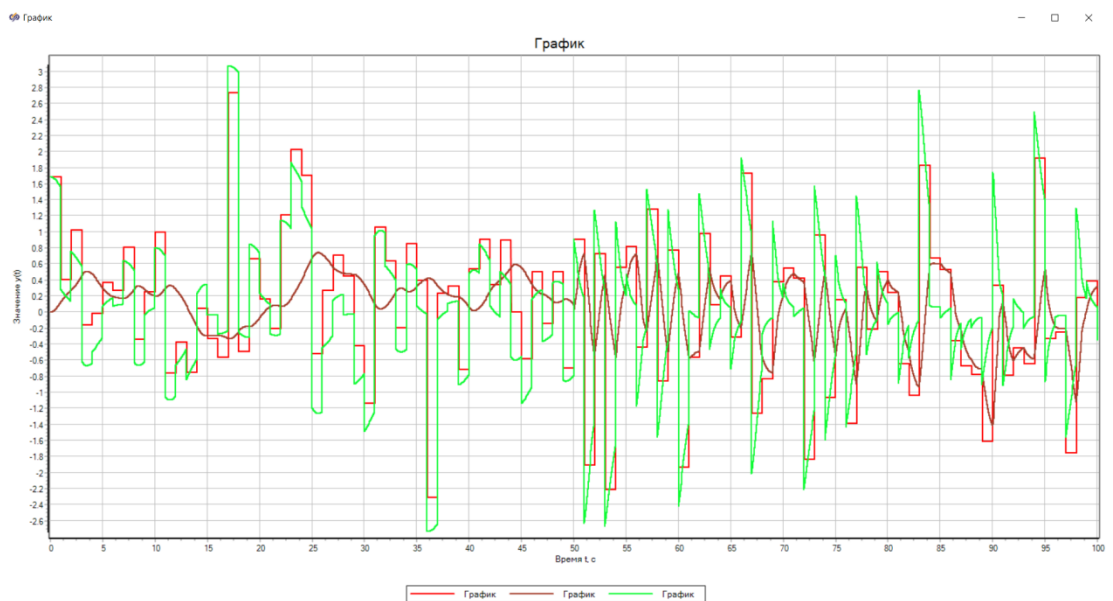


Рис. 8. Переходной процесс нечеткой системы при нормальном шуме

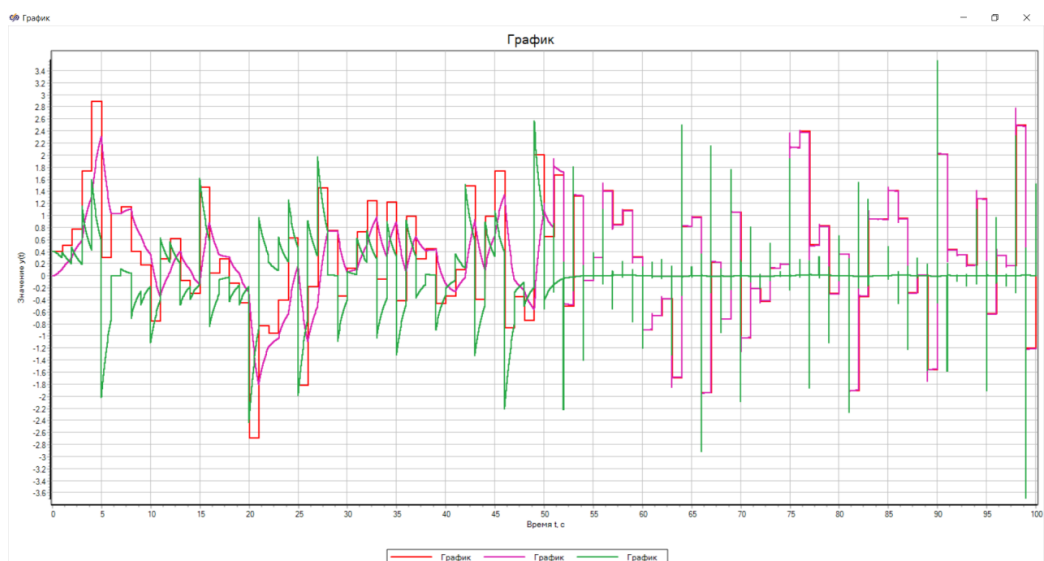


Рис. 9. Переходной процесс классической системы при нормальном шуме

Анализ показателей качества

В результате проведения серии экспериментов на различных входных воздействиях при изменяемых параметрах объекта были получены переходные процессы, представленные на рис. 4. – рис. 9. Далее было проведено исследование качества каждой системы.

Время регулирования t_p можно определить таким образом: по переходному процессу $y(t)$ строится коридор относительно установившегося значения толщиной в два допустимых отклонения, равных 5%. Время регулирования – последняя точка пересечения графика переходного процесса с границей коридора, следовательно, в это время колебания не превышают 5% от установившегося значения. Для более качественной системы необходимо выбрать настройки регулятора, позволяющие обеспечивать наименьшее возможное из значений общего времени регулирования.

Будем определять величину перерегулирования по формуле, приведённой ниже:

$$\sigma = \frac{y_{max} - y_{уст}}{y_{уст}} \cdot 100\%$$

где y_{max} – первое максимальное значение, а $y_{уст}$ – установившееся значение сигнала.

Среднеквадратичная ошибка управления на интервале времени от t_1 до t_N рассчитывалась по формуле:

$$\overline{\varepsilon^2} = \frac{1}{N} \sqrt{\sum_{k=1}^N (r_k - y_k)^2}.$$

Результаты имитационных экспериментов сведены в таблицу 1 с показателями качества управления системой.

Таблица 1

Показатели качества управления систем

Вариант системы управления	Сигнал уставки	Период измерения			
		0-49 с		50-100 с	
Критерий: перерегулирование, время установления					
ПД регулятор	Меандр	0%	6с	12%	2с
НЛР ПД регулятор		0%	8с	0%	3с
Критерий: среднеквадратичная ошибка					
ПД регулятор	Равномерный шум	0,477		0,253	
НЛР ПД регулятор		0,299		0,201	
Критерий: среднеквадратичная ошибка					
ПД регулятор	Нормальный шум	0,943		0,766	
НЛР ПД регулятор		0,805		0,715	

Исходя из приведённых выше показателей качества исследованных систем управления можно сделать выводы:

1. При сигнале уставки «меандр» до изменения параметров обе системы работают без перерегулирования, после изменения параметров объекта система на основе ПД-регулятора имеет перерегулирование 12%, однако время регулирования обеих систем после изменения параметров гораздо ниже.

2. При случайных сигналах уставки после изменения параметров объекта управления среднеквадратичная ошибка ниже, чем до изменения параметров.

3. При случайных сигналах уставки среднеквадратичная ошибка с нечетким ПД-регулятором 1-го рода ниже, чем у систем с традиционным ПД-регулятором.

4. Среднеквадратичные ошибки исследованных систем при уставке «равномерный шум» ниже, чем при уставке «нормальный шум».

Заключение

Разработана модель системы управления в пакете SimInTech и опробованы два алгоритма управления: с традиционным ПД регулятором и нечетким ПД регулятором 1-го рода. На уставке вида «меандр» оба регулятора сопоставимы по качеству управления. На случайных уставках лучшее качество управления обеспечивал нечеткий регулятор. Таким образом, можно отметить перспективность применения нечеткого логического ПД-регулятора 1-го рода для управления нестационарным объектом и целесообразность дальнейших исследований систем управления на основе нечеткой логики.

Библиографический список

1. Надеждин И.С., Горюнов А.Г., Маненти Ф. Системы управления нестационарным объектом на основе МРС-регулятора и ПИД-регулятора с нечеткой логикой, УБС, 2018, выпуск 75, С.50–75
2. Котова Т.С. Построение системы управления на основе нечеткого регулятора первого рода // Тезисы докладов XXVI Международной научно-технической конференции «Радиоэлектроника, электротехника и энергетика». – М.: Изд-во МЭИ, 2020. – С. 253.
3. Анисимов Д.Н. Нечеткие алгоритмы управления. Учебное пособие. – М.: Издательство МЭИ, 2004.
4. Анисимов Д.Н. Использование нечеткой логики в системах автоматического управления // Приборы и системы. Управление, контроль, диагностика. – 2001. – № 8. – С.39-42.
5. Май Тхе Ань. Разработка алгоритмов настройки нечетких логических регуляторов на основе исследования их динамических характеристик: дисс. канд. техн. наук.– М. 2018.
6. Борисов В.В., Круглов В.В., Федулов А.С. Нечеткие модели и сети. – М.: Горячая линия –Телеком, 2012.
7. Коновалов В.В. Повышение производительности токарной обработки колец подшипников путём определения рациональных режимов резания на основе идентификации передаточной функции динамической системы по виброакустическим колебаниям: дисс. канд. техн. наук.– Саратов, 2012.
8. Справочная система Sim In Tech [Электронный ресурс] URL:<https://help.simintech.ru>

T.S. Kotova ¹, V.L. Eliseev ^{1,2}

INVESTIGATION OF FUZZY LOGIC CONTROL SYSTEM WITH NON-STATIONARY OBJECT

¹ *National Research University “Moscow Power Engineering Institute”*,

² *JSC Infotecs*

The problem of synthesis of a feedback control system for an object, the properties of which change stepwise during operation, is considered. A control system with regulators of two types - classical PD and fuzzy logic PD is investigated. The results of simulation modeling in the SimInTech software of a control system with various setpoint signals and control quality criteria are presented. The analysis is carried out and recommendations for more efficient use of fuzzy logic controllers for the control of non-stationary objects are developed.

Keywords: control system, fuzzy logic controller, non-stationary object, PD-controller.

Е.А. Милюкова¹, В.Л. Елисеев^{1,2}

ИССЛЕДОВАНИЕ НЕЙРОСЕТЕВОЙ ОБФУСКАЦИИ ДЛЯ ОБРАБОТКИ КОНФИДЕНЦИАЛЬНЫХ ДАННЫХ

¹Национальный исследовательский университет «МЭИ»

²ОАО «ИнфоТеКС»

Исследуется возможность обработки конфиденциальных данных в зашифрованном виде с помощью нейронных сетей прямого распространения. Описывается процедура создания обучающих данных и обучения нейронных сетей, обфусцирующих алгоритм вычислений. Проводится сравнение архитектуры нейронных сетей и процесса обучения при обработке данных в зашифрованном и незашифрованном виде.

Ключевые слова: искусственные нейронные сети, обфускация, многослойный персептрон, гомоморфное шифрование.

Введение

Современное развитие информационных технологий и вычислительной техники приводит к изменению подхода к обработке данных. Многие компании отказываются от владения масштабной компьютерной инфраструктурой в пользу применения облачных сервисов, которые предоставляют в пользование вычислительные ресурсы, а данные обрабатываются удаленно в крупных центрах. Это позволяет компаниям снизить расходы на владение собственной инфраструктурой. Однако существует проблема, заключающаяся в том, что, хотя алгоритм и данные передаются по защищенным каналам, на стороне облачного сервиса данные и алгоритм обрабатываются в открытом виде. Между тем обрабатываемые данные могут иметь конфиденциальный характер, а алгоритм может являться коммерческой тайной или интеллектуальной собственностью. В таком случае делегация вычислений в недоверенную облачную среду является недопустимой. Для решения этой проблемы может применяться процедура обфускации вычислительного алгоритма, однако это не решает задачу обработки конфиденциальных данных.

Одним из подходов к обфускации является применение нейронных сетей, в котором нейронные сети выступают в качестве математического инструмента, позволяющего реализовать различные функциональные зависимости между входными и выходными данными для элементов алгоритма. При этом предлагается использовать свойство черного ящика, изначально присущего достаточно сложной обученной нейросети. Для защиты конфиденциальных данных требуется их обработка в зашифрованном виде, для чего также подходит аппарат искусственных нейронных сетей. В случае, если защищенный обфускацией алгоритм может обрабатывать зашифрованные данные, возможна передача вычислений в недоверенную среду.

В данной работе исследуется возможность построения нейронной сети, позволяющей эквивалентно заменить вычисление значений некоторой функции в явном виде над незашифрованными данными на нейросетевые вычисления над зашифрованными данными. Рассматриваются три функции различных классов, производится подготовка открытых и зашифрованных обучающих данных, обучение нейронных сетей эквивалентному вычислению значений функций. Подбираются параметры структуры нейронных сетей и проводится их сравнение при обработке данных в открытом и зашифрованном виде.

Теоретический обзор

Обфускация – это один из методов защиты программного кода, который позволяет усложнить процесс понимания работы кода защищаемого программного продукта. Суть процесса обфускации заключается в том, чтобы запутать программный код и устранить

большинство логических связей в нем, то есть трансформировать его так, чтобы он был очень труден для изучения и модификации посторонними лицами (будь то взломщики, или программисты, которые собираются узнать уникальный алгоритм работы защищаемой программы).

В работе [1] было введено строгое понятие стойкой обфускации, основанное на неразличимости, подобной одноименному свойству для криптографических алгоритмов. Неформально, обфускатор O – это транслятор, принимающий на вход программу P и производящий программу $O(P)$, обладающую свойством функциональности и свойством виртуального черного ящика. Свойство функциональности означает, что $O(P)$ производит то же действие, что и P . Свойство виртуального черного ящика означает, что нет более эффективного способа для изучения обфусцированного алгоритма, чем его многократный запуск. Таким образом, можно сказать, что программа, выдаваемая обфускатором, должна давать не больше информации, чем «черный ящик». Было доказано, что нейронная сеть может выступать обфускатором неразличимости [5].

В работе [2] был предложен и апробирован подход к обфускации алгоритмов с использованием искусственных нейронных сетей. Он заключается в представлении исходного алгоритма в виде блок-схемы и замены вычисления значений функций в явном виде на вычисления с использованием нейронных сетей. Предполагалось, что обученная нейронная сеть представляет собой «черный ящик», и нет более эффективного способа понимания логики её работы, чем полный перебор входных аргументов. Слабым местом данного подхода является то, что несмотря на стойкую обфускацию алгоритма, данные обрабатываются в явном виде, и у потенциального злоумышленника по-прежнему остается возможность полного перебора входных значений нейронной сети и анализа реализуемой ею функции.

Для устранения указанного недостатка было предложено обрабатывать обфусцированным алгоритмом зашифрованные данные [6]. В работе [3] была предложена структура системы доверенных облачных вычислений на основе данного подхода. В качестве алгоритма шифрования обрабатываемых данных было предложено использовать таблицы случайных подстановок. Нейронные сети обучаются обрабатывать зашифрованные данные и возвращать результаты в зашифрованном виде, таким образом реализуя гомоморфный подход к шифрованию. При введении шифрования появляется возможность вынести доверенные вычисления с помощью нейронных сетей на удаленный компьютер, в том числе, находящийся в недоверенной среде, где злоумышленник будет иметь возможность исследовать как алгоритм обработки данных, так и сами данные. Подготовка обфусцированного алгоритма, выбор таблиц случайных подстановок, шифрование исходных данных и дешифрование результатов вычислений происходит на доверенной локальной машине.

Метод

Нейросетевая обфускация некоторой функции $f(x)$ с использованием шифрования данных включает в себя:

- создание случайных таблиц подстановки для входных и выходных данных функции, являющихся ключами шифрования;
- шифрование входных данных;
- создание нейронной сети, выполняющей вычисление значения функции над зашифрованными данными и возвращающей результат в зашифрованном виде;
- дешифрование результатов вычисления.

Для простоты, но не снижая общности подхода, далее будем считать, что функция принимает и возвращает целочисленные значения. Тогда можно представить каждый аргумент и значение функции в виде вектора двоичных значений – разрядов двоичного представления. Входы и выходы ИНС в этом случае можно также рассматривать как двоичные значения, причем, выходные нейроны ИНС, имеющие результатом вычислений

вещественные числа *output*, переводятся в двоичные разряды выхода нейронной сети с помощью пороговой функции вида:

$$b = \begin{cases} 0, & \text{output} < 0.5 \\ 1, & \text{output} \geq 0.5 \end{cases}$$

Для создания обучающих данных для нейронной сети требуется ограничить диапазон входных значений аргумента функции на некотором отрезке $[x_{min}; x_{max}]$, который определяется конкретной прикладной задачей. Тогда двоичная разрядность аргумента функции будет равна $\lceil \log_2(x_{max} + 1) \rceil$

Диапазон выходных значений функции определяется самой функцией. Пусть он равен $[y_{min}; y_{max}]$. Тогда, по аналогии, двоичная разрядность значения функции будет равна $\lceil \log_2(y_{max} + 1) \rceil$.

Шифрование входных и выходных данных осуществляется с использованием случайных таблиц подстановки. Шифр подстановки — это метод шифрования, в котором элементы исходного открытого текста заменяются зашифрованным текстом в соответствии с некоторым правилом. Далее таблицей подстановки будет называться взаимно-однозначное соответствие исходных двоичных векторов и зашифрованных двоичных векторов, сгенерированных случайно.

Осуществив нужное количество вычислений исходной функции $f(x)$, будут построены зашифрованная и незашифрованная таблицы пар вход-выход. Имея таблицу пар вход-выход можно выбрать подходящую архитектуру ИНС и обучить её любым алгоритмом с учителем. Целью обучения является нулевая ошибка двоичного выхода нейронной сети. Обучение осуществляется до тех пор, пока среднеквадратическая ошибка на последнем обучаемом слое нейронной сети не станет столь мала, что после применения к выходам нейронов пороговой функции ИНС будет выдавать точное совпадение по всем разрядам для каждого обучающего примера. Поскольку обучение должно осуществляться на полном наборе возможных пар входов-выходов, не требуется осуществлять обобщение, беспокоиться о возможном переобучении и разделять выборку данных на обучающую, тестовую и контрольную.

Таким образом, нейросетевое вычисление над открытыми данными заменяется на последовательность из трех операций: шифрования входных данных, произведение нейросетевых вычислений над ними и дешифрование результата (рис. 1).

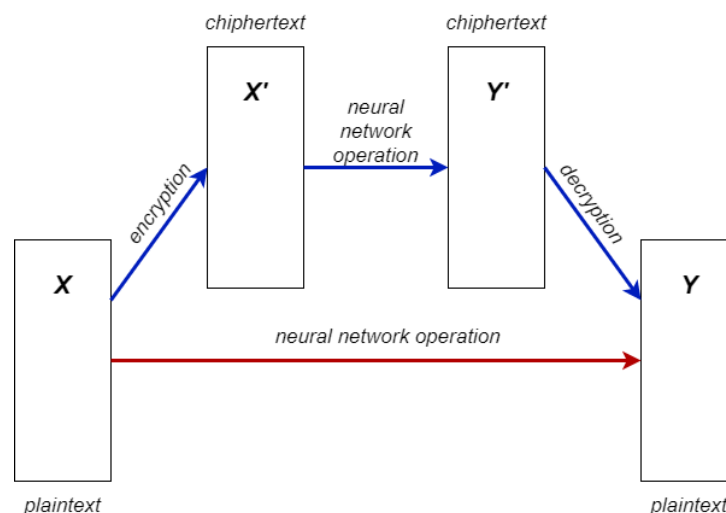


Рис. 1. Схема обработки данных с помощью нейронных сетей.

Для выполнения операций используются модели нейронных сетей прямого распространения (многослойный персептрон). Процесс построения нейронной сети, которая бы решала поставленную задачу, требует решения о выборе количества скрытых слоев и числа нейронов в них, функций активации нейронов, функции потерь, алгоритма обучения и его параметров, числа эпох обучения.

Выбор архитектуры нейронной сети адекватной для конкретной прикладной задачи, заданной обучающим множеством, представляет собой нерешенную проблему. Как правило, архитектура нейронной сети выбирается экспертом и уточняется путем проб и ошибок. Для обоснованного суждения о минимально достаточной сложности нейронной сети, реализующей заданную функцию, применяется следующий эвристический алгоритм [4]:

Шаг 1. В качестве начальной архитектуры выбирается однослойный персептрон с количеством входов и выходов, заданным целевой функцией.

Шаг 2. Если текущая архитектура ИНС в течение заданного фиксированного и достаточно большого количества эпох оказалась обучена с нулевой ошибкой на двоичных выходах, то алгоритм останавливается и ИНС считается результатом.

Шаг 3. В противном случае в нейронную сеть добавляется один скрытый слой, состоящий из одного или нескольких нейронов.

Шаг 4. Если текущая архитектура ИНС в течение заданного фиксированного и достаточно большого количества эпох оказалась обучена с нулевой ошибкой на двоичных выходах, то алгоритм останавливается и ИНС считается результатом.

Шаг 5. В противном случае в скрытый слой ИНС добавляется один или несколько нейронов и происходит переход к шагу 4.

Экспериментальная часть

В качестве функций, над которыми проводится нейросетевая обфускация, выберем следующие функции: ступенчатая функция – $f_1(x)$, сумма двух чисел по модулю 16 – $f_2(x)$ и возведение в квадрат по модулю 32 – $f_3(x)$.

$$f_1(x) = \begin{cases} 0, & 0 \leq x < 18 \\ 5, & 18 \leq x < 30 \\ 6, & 30 \leq x < 50 \\ 4, & x \geq 50 \end{cases}$$

$$x \in \mathbb{Z} \cap [0; 127].$$

$$f_2(x_1, x_2) = (x_1 + x_2) \bmod 16$$

$$x_1, x_2 \in \mathbb{Z} \cap [0; 15].$$

$$f_3(x) = x^2 \bmod 32$$

$$x \in \mathbb{Z} \cap [0; 31].$$

Опишем процесс подготовки данных для функции $f_1(x)$. Для всех целочисленных значений $x \in [0; 127]$ рассчитаны значения функции $y = f(x)$. Далее входные и выходные значения переведены в двоичное представление. Фрагмент данных представлен ниже.

$$x = \begin{bmatrix} \dots \\ 15 \\ 16 \\ 17 \\ 18 \\ 19 \\ 20 \\ 21 \\ 22 \\ 23 \\ \dots \end{bmatrix} \quad y = \begin{bmatrix} \dots \\ 0 \\ 0 \\ 0 \\ 5 \\ 5 \\ 5 \\ 5 \\ 5 \\ 5 \\ \dots \end{bmatrix} \quad x_{bin} = \begin{bmatrix} \dots \\ 0001111 \\ 0010000 \\ 0010001 \\ 0010010 \\ 0010011 \\ 0010100 \\ 0010101 \\ 0010110 \\ 0010111 \\ \dots \end{bmatrix} \quad y_{bin} = \begin{bmatrix} \dots \\ 000 \\ 000 \\ 000 \\ 101 \\ 101 \\ 101 \\ 101 \\ 101 \\ 101 \\ \dots \end{bmatrix}$$

Далее входные и выходные данные зашифровываются с помощью случайной таблицы подстановки. Пример правила, по которому создается таблица подстановки, представлен ниже.

Словарь случайного взаимно-однозначного соответствия для входного аргумента x :

{0: 45, 1: 90, 2: 101, 3: 122, 4: 125, 5: 71, 6: 83, 7: 61, 8: 9, 9: 8, 10: 7, 11: 114, 12: 68, 13: 11, 14: 17, 15: 109, 16: 63, 17: 120, 18: 53, 19: 18, 20: 47, 21: 50, 22: 21, 23: 5, 24: 99, 25: 25, 26: 38, 27: 30, 28: 39, 29: 103, 30: 20, 31: 16, 32: 69, 33: 22, 34: 80, 35: 124, 36: 31, 37: 70, 38: 118, 39: 3, 40: 108, 41: 127, 42: 32, 43: 67, 44: 34, 45: 95, 46: 48, 47: 2, 48: 93, 49: 4, 50: 115, 51: 43, 52: 96, 53: 0, 54: 112, 55: 15, 56: 74, 57: 62, 58: 81, 59: 36, 60: 78, 61: 49, 62: 12, 63: 86, 64: 106, 65: 29, 66: 35, 67: 72, 68: 102, 69: 82, 70: 105, 71: 27, 72: 89, 73: 123, 74: 119, 75: 104, 76: 79, 77: 100, 78: 88, 79: 64, 80: 14, 81: 116, 82: 1, 83: 58, 84: 85, 85: 60, 86: 33, 87: 77, 88: 44, 89: 51, 90: 37, 91: 26, 92: 41, 93: 75, 94: 57, 95: 46, 96: 6, 97: 28, 98: 56, 99: 111, 100: 40, 101: 92, 102: 113, 103: 24, 104: 121, 105: 87, 106: 23, 107: 107, 108: 10, 109: 76, 110: 98, 111: 126, 112: 59, 113: 110, 114: 13, 115: 66, 116: 55, 117: 117, 118: 42, 119: 91, 120: 52, 121: 73, 122: 65, 123: 84, 124: 94, 125: 54, 126: 97, 127: 19}

Словарь случайного взаимно-однозначного соответствия для выхода функции y :

{0: 6, 4: 3, 5: 2, 6: 7}

В соответствии с правилом подстановки входные и выходные данные зашифровываются и переводятся в двоичное представление на семиразрядной сетке. Фрагмент зашифрованных данных представлен ниже.

$$x = \begin{bmatrix} \dots \\ 15 \\ 16 \\ 17 \\ 18 \\ 19 \\ 20 \\ 21 \\ 22 \\ 23 \\ \dots \end{bmatrix} y = \begin{bmatrix} \dots \\ 0 \\ 0 \\ 0 \\ 5 \\ 5 \\ 5 \\ 5 \\ 5 \\ 5 \\ \dots \end{bmatrix} x' = \begin{bmatrix} \dots \\ 109 \\ 63 \\ 120 \\ 53 \\ 18 \\ 47 \\ 50 \\ 21 \\ 5 \\ \dots \end{bmatrix} y' = \begin{bmatrix} \dots \\ 6 \\ 6 \\ 6 \\ 2 \\ 2 \\ 2 \\ 2 \\ 2 \\ 2 \\ \dots \end{bmatrix} x'_{bin} = \begin{bmatrix} \dots \\ 1101101 \\ 0111111 \\ 1111000 \\ 0110101 \\ 0010010 \\ 0101111 \\ 0110010 \\ 0010101 \\ 0000101 \\ \dots \end{bmatrix} y'_{bin} = \begin{bmatrix} \dots \\ 110 \\ 110 \\ 110 \\ 010 \\ 010 \\ 010 \\ 010 \\ 010 \\ 010 \\ \dots \end{bmatrix}$$

На созданных данных независимо обучаются две нейронные сети: одна на незашифрованных данных x_{bin}, y_{bin} , а другая на зашифрованных данных x'_{bin}, y'_{bin} .

Аналогичным образом создаются открытые и зашифрованные обучающие данные для функций f_2 и f_3 , на которых далее обучаются их нейросетевые аналоги. Параметры обученных нейронных сетей и характеристики процесса их обучения представлены в таблице 1.

Обучение производилось алгоритмом стохастического градиентного спуска со скоростью обучения (learning rate) 0.99 и среднеквадратической функцией потерь (meansquared error). Процесс обучения реализован программно на языке Python 3 с использованием библиотек TensorFlow, Keras, NumPy.

Таблица 1

Параметры нейронных сетей и характеристики процесса обучения

	$f_1(\odot)$		$f_2(\odot)$		$f_3(\odot)$	
	Открытые данные	Зашифрованные данные	Открытые данные	Зашифрованные данные	Открытые данные	Зашифрованные данные
Кол-во входных нейронов	7		8		5	
Кол-во выходных нейронов	3		4		5	
Кол-во нейронов в скрытом слое	5	10	15	30	5	10
Кол-во весовых коэффициентов	58	113	199	394	60	115
Кол-во эпох обучения	5000	10000	5000	10000	10000	5000
Время обучения, сек	13.8	28.6	17.2	50.0	8.1	4.0

Анализ результатов

Результаты проведенных экспериментов показывают, что нейронным сетям, обрабатывающим зашифрованные данные, для корректного обучения требуется больше нейронов в скрытом слое и, как следствие, увеличивается число весовых коэффициентов. Тем самым, при обработке зашифрованных данных растет вычислительная сложность операций. Для нейронных сетей, реализующих функции f_1 и f_2 потребовалось больше эпох обучения при обработке зашифрованных данных, что увеличило время обучения примерно в 2 раза. Для функции f_3 наблюдается обратная картина, но следует заметить, что длительность процесса обучения сильно зависит от начальной инициализации весовых коэффициентов сети и при удачной инициализации обучение может проходить быстрее.

Важно отметить, что для вычисления значения функции с помощью обученной нейронной сети, необходимо сначала провести операцию шифрования входного значения, вычислить выходное зашифрованное значение и затем расшифровать результат. Всё это приводит к увеличению вычислительных операций для обработки данных с помощью нейросетевой обфускации исходной функции. Однако введение шифрования очевидно избавляет алгоритм обфускации от недостатка, заключающегося в том, что для понимания функции, реализуемой нейросетью, достаточно сделать полный перебор входных данных. Произведя полный перебор входов для нейросети, обученной на зашифрованных данных, нельзя будет понять исходную функцию.

Заключение

В работе был реализован и опробован подход к нейросетевой обфускации над зашифрованными данными, а также произведен анализ влияния шифрования данных на структурную сложность и время обучения нейронных сетей.

Рассмотренный подход позволит обфусцировать некоторый вычислительный алгоритм, в котором будет использоваться множество различных функций и их композиций. Обфусцированный алгоритм, работающий с зашифрованными данными, может быть передан в недоверенную среду облачного агента. В таком случае возможен обмен зашифрованными данными с облачным агентом без необходимости доверять ему, так как данные будут защищены шифрованием, а алгоритм обфускацией.

Библиографический список

1. Barak B., Goldreich O., Impagliazzo R., Rudich S., Sahai A., Vadhan S., and Yang K. On the (im)possibility of obfuscating programs, in Advances in Cryptology CRYPTO. Springer, 2001.
2. Колпинский С.В., Милюкова Е.А., Елисеев В.Л., Нейросетевая обфускация алгоритма банковского скоринга // Сборник материалов XXV Международной научно-технической конференции «Информационные системы и технологии – 2019», Нижний Новгород, 19 апреля 2019, с.725–732.
3. Милюкова Е. А., Елисеев В. Л., Проектирование системы доверенных облачных вычислений с использованием нейросетевой обфускации // Сборник трудов XXIX Международной научно-технической конференции «Современные технологии в задачах управления, автоматизации и обработки информации», Алушта, 14-20 сентября 2020, с. 166-167.
4. Милюкова Е.А., Елисеев В.Л., Нейросетевая реализация некоторых классов булевых функций // Сборник материалов XXVI Международной научно-технической конференции «Информационные системы и технологии – 2020», Нижний Новгород, 24 - 28 апреля 2020, с.786–797.
5. Елисеев В.Л., Искусственные нейронные сети как механизм обфускации вычислений // Труды всероссийской конференции «XVIII Сибирская научная школа-семинар с международным участием «Компьютерная безопасность и криптография» – SIBECRYPT'19» Томск 9-14 сентября 2019 г., Прикладная дискретная математика. Приложение, 2019, №12, с.165-169.
6. Елисеев В.Л., Нейросетевая обфускация вычислений над зашифрованными данными // Прикладная дискретная математика. Приложение. 2020, №13, С.85-93. DOI: 10.17223/2226308X/13/25

Е.А. Miliukova¹, V.L. Eliseev^{1,2}

RESEARCH OF NEURAL NETWORK OBFUSCATION FOR PROCESSING CONFIDENTIAL DATA

¹ *National Research University “Moscow Power Engineering Institute”*,

² *JSC Infotecs*

The possibility of processing confidential data in encrypted form using neural networks of direct distribution is investigated. The procedure for producing training data and training neural networks that obfuscate the calculation algorithm is described. We compare the architecture of neural networks and the learning process when processing data in encrypted and non-encrypted form.

Keywords: artificial neural networks, obfuscation, multi-layer perceptron, homomorphic encryption.

ОБНАРУЖЕНИЕ АТАК НА АСУ ТП МЕТОДОМ АНОМАЛИЙ С ПРИМЕНЕНИЕМ НЕЙРОННЫХ СЕТЕЙ LSTM

¹Национальный исследовательский университет «МЭИ»

²ОАО «ИнфоТеКС»

В работе рассматривается задача обнаружения сетевых атак на промышленные сети АСУ ТП методом аномалий. Предлагается методика применения рекуррентных нейронных сетей для обнаружения атак как аномалий технологического процесса. Для апробации подхода используются данные, полученные на испытательном стенде по очистке воды Secure Water Treatment (SWaT). Проводится серия экспериментов с обнаружением аномалий в тестовых данных. В качестве детектора аномалий используется искусственная нейронная сеть, содержащая ячейки с обратными связями – Long Short-Term Memory (LSTM). Описывается структура построенного детектора, вычисляются метрики качества обнаружения аномалий, делаются выводы о свойствах детектора и применимости предложенной методики обнаружения атак.

Ключевые слова: обнаружение аномалий, промышленные сети, безопасность АСУ ТП, рекуррентные нейронные сети, LSTM.

Введение

Автоматизированные системы управления технологическими процессами (АСУ ТП) давно стали необходимой частью современного производства. Сменилось несколько поколений аппаратуры средств автоматизации. С развитием технологий АСУ ТП постепенно преобразовались из закрытых управляющих систем в многоуровневые промышленные сети на базе стандартных сетевых протоколов, которые имеют значительное сходство с широко распространенными корпоративными сетями. К сожалению, это касается и уязвимостей, которые связаны с угрозами кибербезопасности. Промышленные сети подвержены заражению вредоносными программами, взлому, выводу из строя программного и аппаратного обеспечения и другим видам внешнего воздействия средствами информационных технологий. Есть известные прецеденты вывода из строя крупных промышленных объектов путем направленной компьютерной атаки. Это ставит вопрос о своевременном обнаружении атак и выработке мер защиты.

Сети АСУ ТП подвержены компьютерным атакам во многом из-за устаревшей архитектуры, не содержащей встроенных средств защиты, а также вследствие подключения прямо или опосредовано к сети Интернет. Переработка архитектуры с целью повышения защищенности в большинстве случаев невозможна без глубокой модернизации технических средств и сопряженных с этим дорогостоящих работ. Поэтому адекватной и экономичной мерой является разработка инструментов для обнаружения атак, что позволит своевременно предотвратить наступление нежелательных последствий, касающихся влияния на собственно технологические процессы производства.

Обзор литературы

Разнообразным технологиям обнаружения компьютерных атак и вторжений посвящено множество работ и обзоров (например, [4]). По-крупному их можно подразделить на три типа: сигнатурные, основанные на спецификациях и основанные на обнаружении аномалий. Принцип обнаружения атак у первых двух технологий заключается в том, что характерные признаки атаки заведомо известны, что позволяет составить правила их выявления. В то же время, наиболее известные атаки на АСУ ТП имели характер целевых, детально разрабатываемых для нанесения ущерба конкретной инфраструктуре. Таким образом, по крайней мере часть вредоносных действий была

уникальной и по определению не могла быть обнаружена методами, построенными на априорном знании об атаке.

Методы, основанные на обнаружении аномалий, характерны тем, что в их основе лежит модель не атаки, а нормального поведения, которое нарушается при атаке на систему. За счет применения этого подхода становится возможным обнаруживать ранее не встречавшиеся атаки как отклонение технологического процесса от модели нормы, построенной на основе данных, описывающих штатную работу системы. Применение такого подхода выглядит перспективным именно для систем АСУ ТП, особенностью которых являются простые и, в основном, циклические сценарии работы. Построение модели нормы для таких процессов методами машинного обучения представляется разрешимой задачей.

Для целей создания систем обнаружения атак и проведения исследований в области защиты промышленных систем исследователями создаются и публикуются общедоступные наборы данных. Для создания датасетов авторы программно моделируют системы АСУ ТП или создают физические лабораторные стенды. В работе [5] проводится подробное сравнение общедоступных наборов данных, созданных для исследования вопросов информационной безопасности в промышленных сетях.

Особый интерес для исследования представляет набор данных SWaT (от англ. Secure Water Treatment – защищенная очистка воды), данные для которого были собраны на испытательном стенде, реализующем технологический процесс очистки воды [9]. Стенд состоит из различных технических устройств для очистки воды, контроллеров и SCADA-системы. В течение 11 дней собирался сетевой трафик, а также записывалась информация с датчиков, исполнительных механизмов, сигналов ввода-вывода ПЛК: 7 дней в режиме нормальной работы и 4 дня при реализации заранее подготовленных атак. В общей сложности были реализованы 36 сценариев атак, связанных с физическими сигналами и сетевым трафиком. В работе [3] приводится подробное описание стенда и реализованных атак: описывается время атаки, узел системы, на который она направлена, состояние узла до атаки и само атакующее воздействие, а также ожидаемый и возникший результат воздействия. Так как данные собираются и на физическом, и на сетевом уровне, то возможно применение этого набора данных для исследования различных подходов интеллектуального анализа данных и машинного обучения для обнаружения атак в промышленных сетях, в частности, анализа аномалий в физических сигналах. Актуальность модели SWaT подтверждается поступающими новостями о взломах функционирующих водоочистных сооружений [8].

Применение метода аномалий для обнаружения сетевых атак в промышленных системах управления основано на том, что сценарии работы одного датчика, исполнительного механизма или контроллера обычно довольно просты и повторяются циклически. Для реализации детектора аномалий может использоваться нейросетевой подход к машинному обучению. Применение рекуррентных нейронных сетей в сочетании со статистическими методами обработки данных позволяет строить качественные детекторы аномалий в промышленных сетях [6, 7]. Однако в докладе [7] даются лишь общие выкладки и не представляется информации о конкретной архитектуре применяемого детектора аномалий и принципе его обучения.

Рекуррентные нейронные сети, то есть сети с обратными связями, разработаны специально для обработки последовательностей. За счет наличия обратных связей они обладают свойством памяти и могут выявлять скрытые зависимости, распределенные во времени [2]. Способность рекуррентных нейронных сетей к обобщению и выявлению скрытых зависимостей между входными и выходными данными, в совокупности со свойством памяти приводит к идее об их применении при построении модели нормы технологического процесса

В работе [1] было описано устройство рекуррентных нейронных сетей LSTM и рассматривалось их применение для выявления аномалий на модельных примерах. Для

обучения нейронной сети моделировался сигнал без аномалий, а сигнал для тестирования отличался от обучающего инъекцией аномалий в нескольких местах. Нейронная сеть обучалась воспроизводить нормальный сигнал с минимальной ошибкой. Когда на вход обученной нейросети подавался тестовый сигнал с аномалиями, то в местах инъекций модель выдавала большую ошибку воспроизведения. Было показано, что использование сигнала ошибки воспроизведения позволяет определять наличие аномалий и места их инъекций. Основываясь на полученных в [1] результатах было предложено применить аналогичный подход для обнаружения атак, воспользовавшись набором данных SWaT [3].

Метод

Целью исследования является проверка гипотезы: применение нейросетевой модели нормы технологического процесса позволяет обнаруживать аномалии, вызванные атаками на АСУ ТП. В качестве обучающих данных для построения нейросетевой модели нормы процесса требуется достаточно длинная запись процесса, которая бы включала в себя отработку всех штатных режимов, но при этом не содержащая аномалий. В качестве тестовых данных для проверки работы детектора аномалий требуется запись процесса, в которую в явном виде были бы внесены отклонения от штатного режима функционирования.

Обучение нейросетевой модели будет проводиться с учителем, а это значит, что для обучения и тестирования требуются не только входные, но и целевые выходные данные. В данном случае модель будет обучаться воспроизводить на выходе то же значение, что и подается на вход. Формально, для входной (как обучающей, так и тестовой) последовательности

$$X = x_1, \dots, x_t, \dots, x_T,$$

где T – количество отсчетов в ней, сопоставляется целевая последовательность

$$Y = y_1, \dots, y_t, \dots, y_T,$$

такая, что

$$y_t = x_t, \quad t = 1 \dots T.$$

За счет наличия обратных связей и специального устройства LSTM, модель не только обучается передавать на выход то же значение, что и поступило на вход, но и сохраняет в себе скрытые состояния и сможет моделировать норму процесса.

Ошибка на шаге t будет определяться как

$$e_t = |y_t - \hat{y}_t|, \quad \hat{y}_t = model(x_t), \quad t = 1 \dots T,$$

где y_t – целевое значение, а \hat{y}_t – значение, полученное с выхода модели при подаче на вход значения x_t .

Качественным результатом обучения модели будет являться то, что она сможет воспроизводить обучающий сигнал с некоторой малой ошибкой, а при подаче на вход модели тестового сигнала с аномалиями ожидается большая ошибка в тех местах сигнала, где включены аномалии.

Для количественной оценки качества детектора аномалий требуется выработать решающее правило, которое по значению ошибки на тестовом сигнале e_t^{test} на шаге t будет определять, является данный отсчет x_t^{test} аномальным или нормальным и присваивать данному отсчету метку *predicted label_t*. В качестве решающего правила выбирается сравнение ошибки на тестовом сигнале с пороговым значением *threshold*:

$$predicted\ label_t = \begin{cases} 'anomal', & e_t^{test} > threshold \\ 'normal', & e_t^{test} \leq threshold \end{cases}$$

Пороговое значение ошибки *threshold* определяется как максимальное значение ошибки на обучающем сигнале:

$$threshold = \max_{t=1 \dots T_{train}} e_t^{train},$$

где T_{train} – количество отсчетов в обучающем сигнале.

После того, как для каждого отсчета в тестовом сигнале будет назначена метка, можно считать, что процесс обнаружения аномалий завершен и для детектора можно вычислить метрики качества обнаружения аномалий, а именно количество истинно положительных (true positive, TP), истинно отрицательных (true negative, TN), ложноположительных (false positive, FP) и ложноотрицательных (false negative, FN) срабатываний. Для этого предполагается, что для тестового сигнала заранее известно положение аномалий и каждому отсчету x_t^{test} априори присвоена истинная метка *true label_t*. В таком случае определим

$$\begin{aligned} TP &= N\{predicted\ label_t = 'anomal' \ \& \ true\ label_t = 'anomal'\} \\ TN &= N\{predicted\ label_t = 'normal' \ \& \ true\ label_t = 'normal'\} \\ FP &= N\{predicted\ label_t = 'anomal' \ \& \ true\ label_t = 'normal'\} \\ FN &= N\{predicted\ label_t = 'normal' \ \& \ true\ label_t = 'anomal'\} \\ & \quad t = 1 \dots T_{test}, \end{aligned}$$

где T_{test} – количество отсчетов в тестовом сигнале, а $N\{\cdot\}$ – количество выполнений условия.

Реализация метода и эксперименты

Для проведения экспериментов из набора данных SWaT были выбраны три сигнала, подверженных атакам. Это сигналы LIT101 (сигнал с датчика уровня воды на первой стадии процесса очистки), DRIT301 (сигнал с датчика перепада давления на третьей стадии процесса очистки) и сигнал LIT301 (сигнал с датчика уровня воды на третьей стадии процесса очистки). В качестве обучающих данных были выбраны данные за период 3 суток с 10:00:00 25.12.2015 по 09:59:59 28.12.2015. В качестве тестовых данных с аномалиями были выбраны данные за период 2 суток с 10:00:00 28.12.2015 по 09:59:59 30.12.2015. За выбранный период на тестовые сигналы было проведено 6 атак: по 2 на каждый сигнал. Графики обучающих и тестовых сигналов с выделенными местами атак представлены на рис. 1. В табл. 1 приведены сведения о проведенных атаках на сигналы в выбранный промежуток времени.

Таблица 1

Описание проведенных атак

Точка атаки	Действие	Дата	Время начала	Время окончания
LIT101	Увеличение на 1 каждую секунду	28.12.2015	11:20:09	11:28:02
LIT101	Установка значения 700	29.12.2015	18:29:59	18:41:40
DRIT301	Установлено значение 45	28.12.2015	13:09:45	13:25:54
DRIT301	Установка значения 45	30.12.2015	01:42:07	01:53:30
LIT301	Установка значения 1200	28.12.2015	12:08:05	12:15:12
LIT301	Уменьшение на 1 каждую секунду	29.12.2015	11:57:04	12:01:44

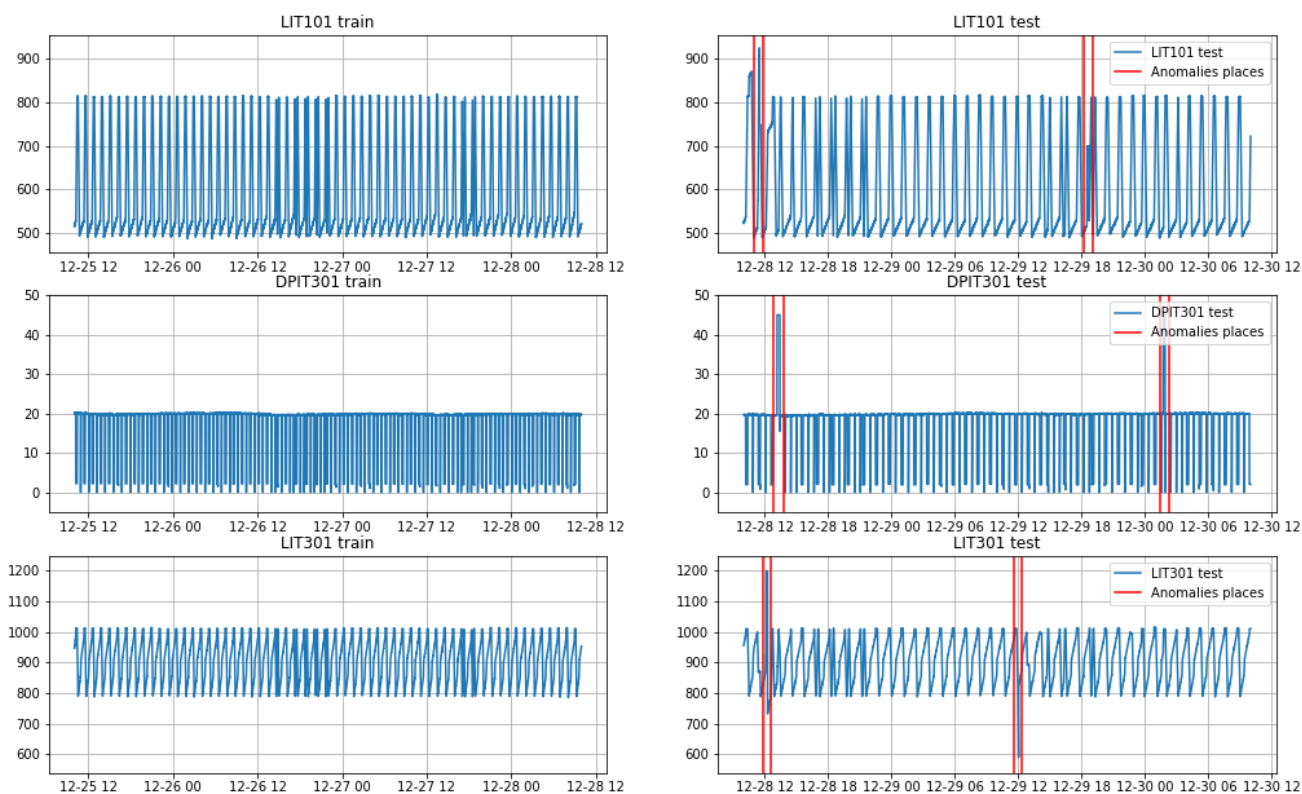


Рис. 1. Графики обучающих и тестовых сигналов с выделенными аномалиями.

Перед построением детекторов и обучением нейросетевых моделей требуется предварительная обработка данных. Данные в датасете SWaT имеют период дискретизации в 1 секунду. В таком случае количество отсчетов в обучающих и тестовых сигналах равно 259 200 и 172 800 соответственно. Для снижения количества отсчетов было произведено прореживание сигналов с периодом в 30 секунд. Таким образом удалось снизить количество отсчетов до 8640 и 5760 соответственно без потери качественных свойств сигналов. Далее было произведено линейное приведение обучающих сигналов к масштабу $[min_{train}; max_{train}] \rightarrow [0; 1]$, а тестовые данные были преобразованы с теми же масштабирующими коэффициентами, что и обучающие.

После проведения предварительной обработки данных для каждого из трех выбранных сигналов были проведены действия согласно описанному выше методу и построено три независимых детектора аномалий. Нейронные сети, используемые в детекторах, имеют структуру, представленную на рис. 2. В качестве нейронов в первом скрытом слое используются рекуррентные LSTM нейроны, во втором скрытом слое и в выходном слое – искусственные нейроны с гиперболическим тангенсом в качестве функции активации. Параметры обучения: функция потерь – среднеквадратическая, алгоритм обучения – стохастический градиентный спуск (SGD) со скоростью обучения (learning rate) 0.01, количество эпох – 50. Детекторы были реализованы программно на языке Python 3 с использованием прикладных библиотек Keras, TensorFlow, NumPy, Pandas, Matplotlib [10-14].

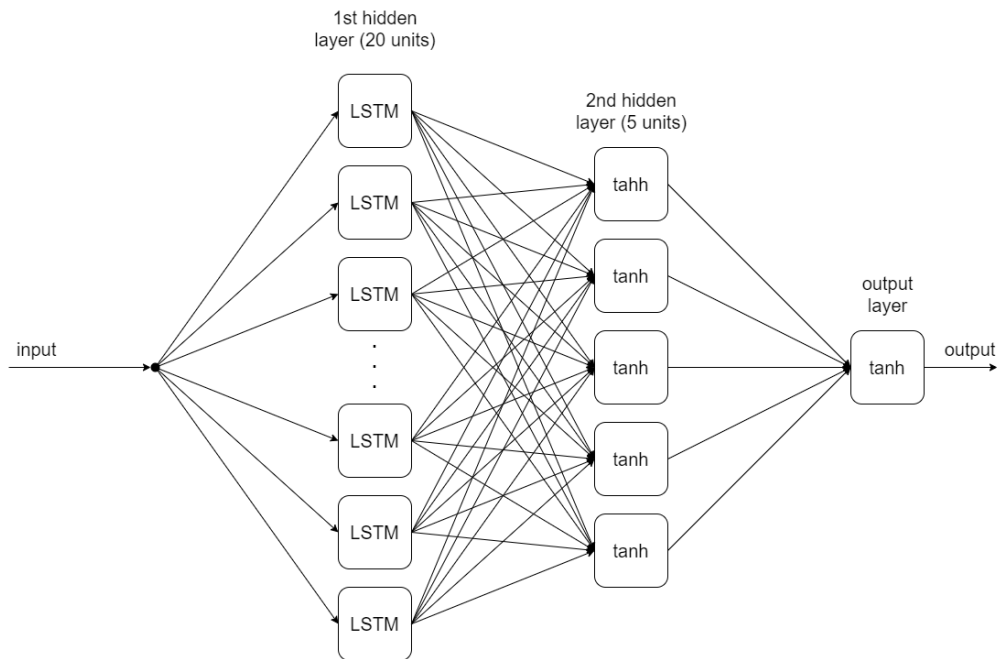


Рис. 2. Структура нейросетевой модели.

В качестве результатов экспериментов на рис. 3-5 приводятся графики тестовых сигналов, совмещенных с сигналами выхода нейросетевой модели и графики ошибки на тестовых данных с порогом принятия решения для трех сигналов: LIT101, DPIT101 и LIT301. Красными линиями обозначены участки атак. На рис. 6 представлены метрики качества обнаружения аномалий.

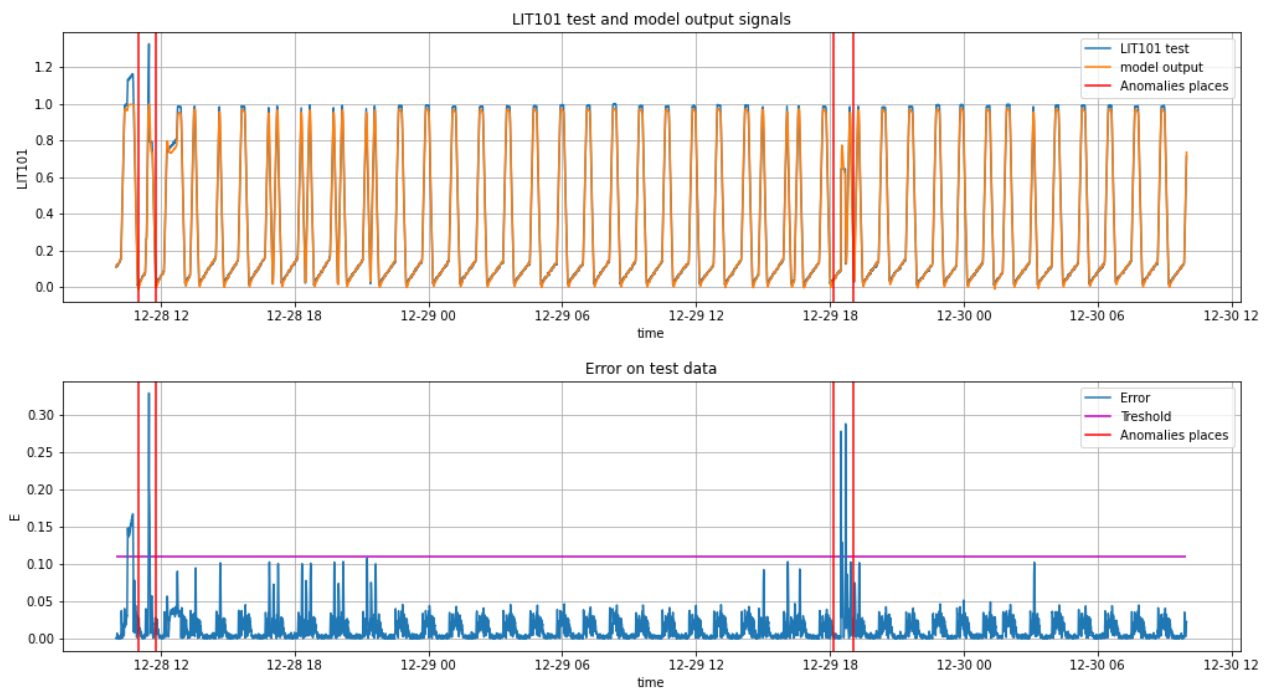


Рис. 3. График тестового сигнала LIT101, сигнала выхода модели и график ошибки.

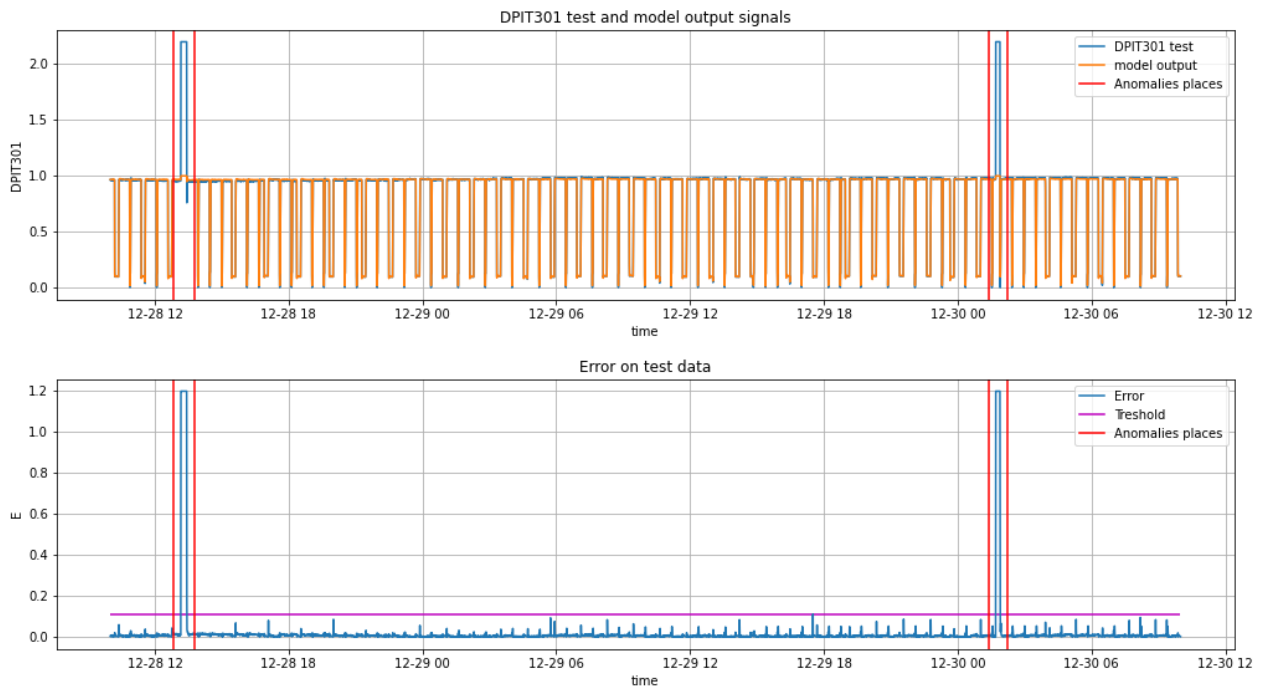


Рис. 4. График тестового сигнала DPIT301, сигнала выхода модели и график ошибки.

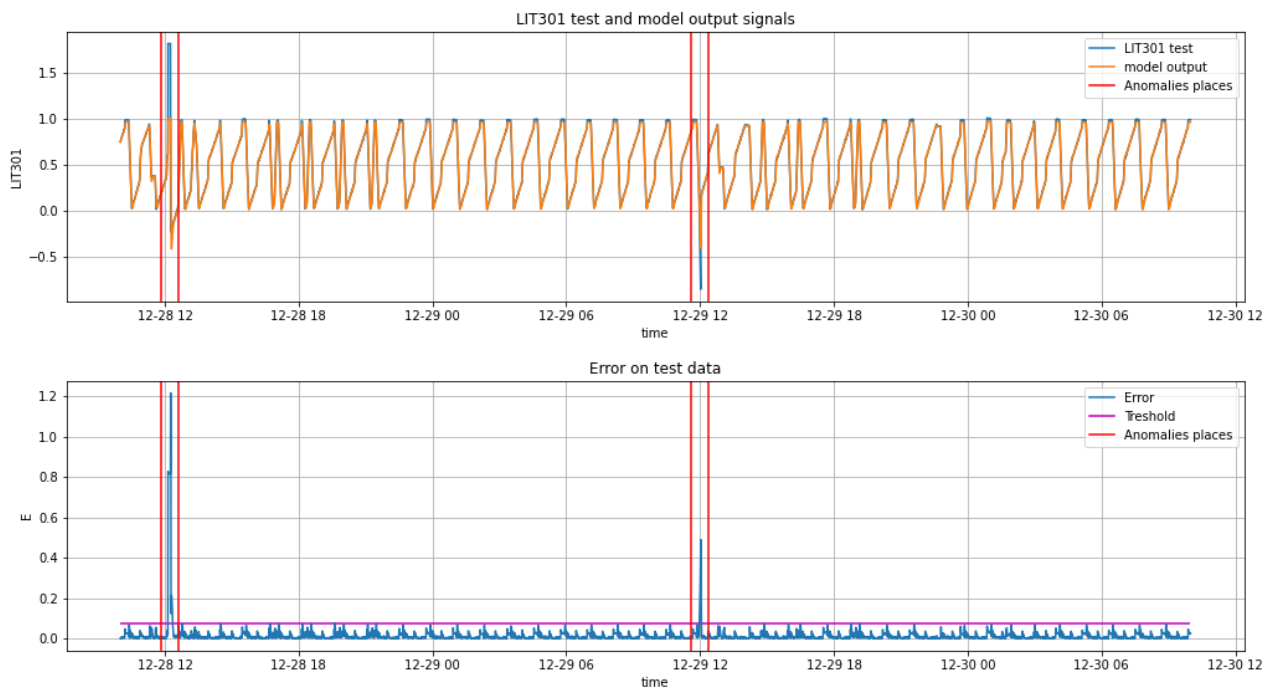


Рис. 5. График тестового сигнала LIT301, сигнала выхода модели и график ошибки.

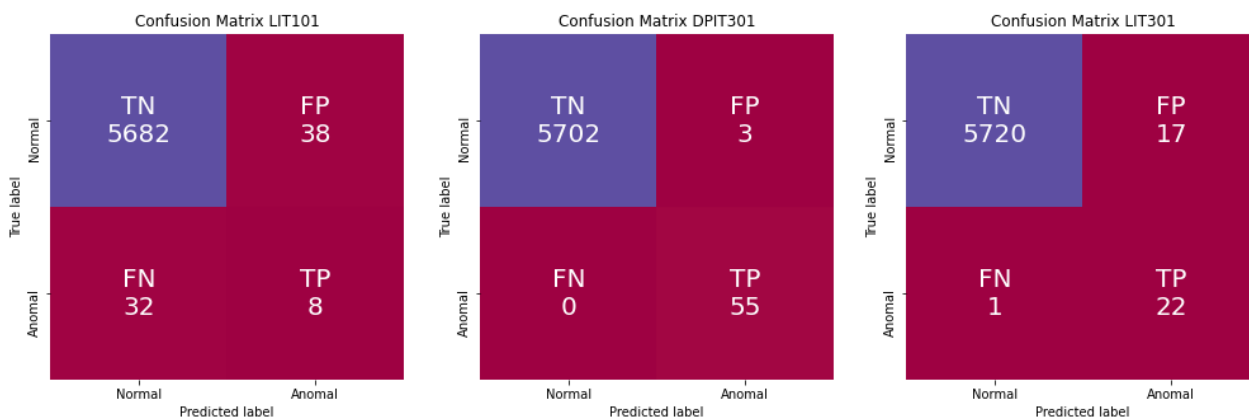


Рис. 6. Метрики качества обнаружения аномалий.

Анализ результатов экспериментов

Результаты проведенных экспериментов показывают, что все созданные детекторы аномалий удовлетворяют качественным свойствам и во всех экспериментах проявляется их ожидаемое поведение. Сигналы ошибки воспроизведения тестовых сигналов оказываются больше порогового значения в местах включения аномалий, и ниже порогового значения вне мест включения аномалий. Таким образом, гипотеза о том, что применение нейросетевой модели нормы технологического процесса позволяет обнаруживать аномалии, вызванные атаками на АСУ ТП, можно считать подтвержденной.

Обращаясь к количественным результатам экспериментов – метрикам качества обнаружения аномалий, стоит заметить, что применительно к безопасности АСУ ТП, ложноотрицательные срабатывания детектора несут более серьезный характер, чем ложноположительные, для системы обнаружения атак желательно свести их количество к минимуму. Ложноотрицательное срабатывание означает пропуск детектором аномалии, что может привести к нарушению хода процесса выводу из строя оборудования. Большое количество ложноположительных срабатываний хоть и говорит о недостаточном качестве детектора, но не приводит к угрозе оборудованию.

Рассмотрим метрики качества обнаружения аномалий в проведенных экспериментах. Для сигнала DPIT301 имеем ни одного ложноотрицательного и 3 ложноположительных срабатывания, что говорит о высоком качестве детектирования аномалий для этого сигнала. В этот сигнал были внесены существенные аномалии, что позволило обнаружить их с высокой точностью. Для сигнала LIT301 имеем 1 ложноотрицательное и 17 ложноположительных срабатываний, и для данного сигнала также можно говорить о высокой точности детектирования. Для сигнала LIT101 метрики качества показывают, что количество ложных срабатываний в сумме на порядок больше, чем истинно положительных, хотя визуально по графику ошибки на этом сигнале видно, что ошибка выше порога в местах аномалий и ниже порога вне их. Этот результат говорит о несовершенстве метода выработки решающего правила для определения аномалий на основании сигнала ошибки воспроизведения и дает направление для дальнейших исследований в этой области.

Говоря в целом о предложенной методике обнаружения атак на промышленные системы как аномалий в технологическом процессе, можно считать ее применимой, особенно в совокупности с другими методами обнаружения атак.

Заключение

В данной работе был предложен метод использования рекуррентных нейронных сетей LSTM для обнаружения атак на промышленные системы, основанный на анализе аномалий. Проведена серия экспериментов с данными, полученными на испытательном стенде по очистке воды SWaT. Достигнутые результаты в части построения детекторов аномалий и метрик их качества на наборе данных SWaT позволяют рассматривать технологию как перспективную для применения в системах обнаружения вторжений.

Библиографический список

1. Колпинский С.В., Елисеев В.Л. Распознавание новизны в потоке битов с помощью LSTM // Сборник материалов XXVI Международной научно-технической конференции «Информационные системы и технологии – 2020». – Нижний Новгород, 24 - 28 апреля 2020. – С. 771–779.
2. Николенко С., Кадурин А., Архангельская Е. Глубокое обучение. – СПб.: Питер, 2019. – 480 с.: ил. – (Серия «Библиотека программиста»). ISBN 978-5-496-02536-2.
3. Goh J., Adepu S., Junejo K.N., Mathur A. A Dataset to Support Research in the Design of Secure Water Treatment Systems. – Singapore, 2016.
4. Hindy H. et al., A Taxonomy of Network Threats and the Effect of Current Datasets on Intrusion Detection Systems in IEEE Access, vol.8, pp.104650-104675, 2020, doi: 10.1109/ACCESS.2020.3000179.
5. Choi S., Yun J.-H., Kim S.-K. A Comparison of ICS Datasets for Security Research based on Attack Paths. – Daejeon, Republic of Korea, 2018.
6. Водоочистительный полигон SWaT [Электронный ресурс]: Kaspersky Machine Learning for Anomaly Detection – примеры применения. URL: <https://mlad.kaspersky.ru/vodoochistitelnyj-poligon-swat/> (дата обращения 10.12.2020).
7. Искусственный интеллект для мониторинга операционных процессов. Существенное расширение возможностей оператора [Электронный ресурс]: Архив докладов Национального нефтегазового форума – 2019. URL: <http://oilandgasforum.ru/data/files/nnf2019/lavrentiev.pdf> (дата обращения 10.12.2020).
8. Киберпреступники взломали водоочистные сооружения в Олдсмаре, Флорида [Электронный ресурс]: URL: <https://www.anti-malware.ru/news/2021-02-09-111332/34935> (дата обращения 05.02.2021).
9. Secure Water Treatment testbed [Электронный ресурс]. URL: <https://itrust.sutd.edu.sg/testbeds/secure-water-treatment-swat/> (дата обращения 15.11.2020).
10. Keras documentation: Developer guides [Электронный ресурс]. URL: <https://keras.io/guides/> (дата обращения 01.02.2021).
11. TensorFlow Core API documentation [Электронный ресурс]. URL: https://www.tensorflow.org/api_docs/python/tf/ (дата обращения 01.02.2021).
12. NumPy documentation [Электронный ресурс]. URL: <https://numpy.org/doc/> (дата обращения 01.02.2021).
13. Pandas documentation [Электронный ресурс]. URL: <https://pandas.pydata.org/docs/> (дата обращения 01.02.2021).
14. Matplotlib documentation [Электронный ресурс]. URL: <https://matplotlib.org/3.3.3/contents.html> (дата обращения 01.02.2021).

S.V. Kolpinskiy ¹, V.L. Eliseev ^{1,2}

DETECTION OF ATTACKS ON THE INDUSTRIAL AUTOMATED CONTROL SYSTEM BY THE METHOD OF ANOMALIES USING LSTM NEURAL NETWORKS

¹ National Research University “Moscow Power Engineering Institute”,

² JSC Infotecs

The paper considers the problem of detecting network attacks on industrial automated control system networks by the method of anomalies. A method of using recurrent neural networks to detect attacks as anomalies of the technological process is proposed. To test the approach, the data obtained at the Secure Water Treatment (SWaT) test bench is used. A series of experiments is being conducted to detect anomalies in the test data. An artificial neural network containing cells with feedback loops – Long Short-Term Memory (LSTM) – is used as an anomaly detector. The structure of the constructed detector is described, machine learning quality metrics are calculated, conclusions about the properties of the detector and the applicability of the proposed method of detecting attacks are drawn.

Keywords: anomaly detection, industrial networks, industrial automated control system security, recurrent neural networks, LSTM.

ПРИБЛИЖЕННЫЙ АЛГОРИТМ РЕШЕНИЯ КВАДРАТИЧНОЙ ЗАДАЧИ О НАЗНАЧЕНИЯХ

Нижегородский государственный университет им. Н. И. Лобачевского

В работе рассмотрена проблема отображения параллельной программы на вычислительную систему с целью оптимизации времени выполнения за счет сокращения времени пересылки данных. Проблема сформулирована в виде квадратичной задачи о назначениях. Разработана и реализована параллельная приближенная модификация метода ветвей.

Ключевые слова: квадратичная задача о назначениях, метод ветвей и границ.

Введение

Вычислительная система является дорогостоящим ресурсом. Для его эффективного использования необходимо оптимизировать время работы программ. Время выполнения может зависеть от алгоритма, реализации, количества вычислителей и других факторов.

Существует множество алгоритмов, которые решаются за реальное время только будучи параллельно распределенными. Для этого используются вычислительные системы с несколькими вычислительными узлами. При этом между ними происходит обмен данными.

Для эффективного использования ресурсов в таких системах требуется решать задачи, связанные с оптимизацией времени работы программ, которые на этой системе выполняются. При этом одной из возникающих проблем является проблема отображения параллельной программы на вычислительную систему [1]. Сформулируем её в виде квадратичной задачи о назначениях, а решение будем искать приближенной модификацией метода ветвей и границ.

Содержательное описание и постановка задачи

Имеется вычислительная система с несколькими узлами, и с определенной стоимостью передачи данных между ними. Есть некоторая задача, параллельный алгоритм которой будем считать уже реализованным. Её подзадачи могут коммуницировать между собой. Надо распределить их по узлам так, чтобы минимизировать стоимость пересылки данных.

Рассмотрим пример параллельной программы из четырех подпрограмм. На рис. 1 представлено её схематичное представление.

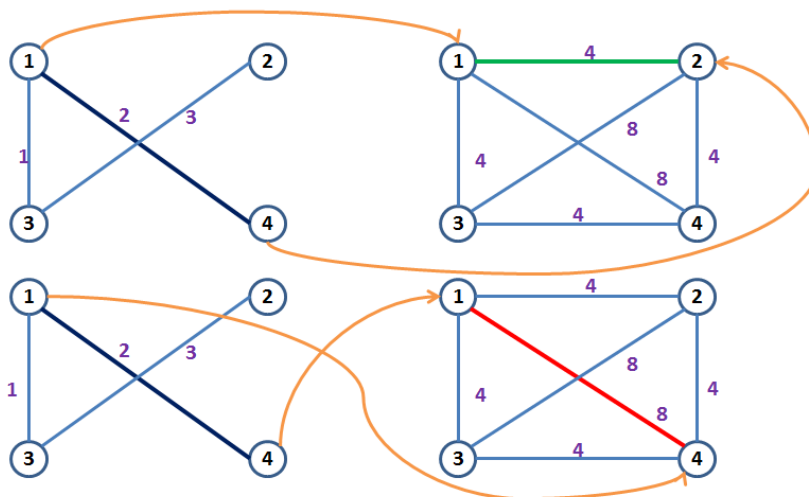


Рис. 1. Схематичное представление параллельной программы

Рассмотрим варианты возможного назначения подзадач по узлам. Назначим подзадачи первым способом: $1 \rightarrow 1, 2 \rightarrow 4, 3 \rightarrow 3, 4 \rightarrow 2$. В этом случае значение критерия равно $(2 * 4 + 1 * 4 + 3 * 4) * 2 = 48$. Назначим другим способом: $1 \rightarrow 4, 2 \rightarrow 3, 3 \rightarrow 2, 4 \rightarrow 1$. Критерий равен $(2 * 8 + 1 * 4 + 3 * 8) * 2 = 88$. При назначении задач первым способом значения критерия меньше, чем при назначении вторым.

Таким образом, назначив задачи по вычислительным узлам оптимальным способом можно сократить время на пересылку данных.

Параллельную программу будем моделировать с помощью квадратной матрицы. Имеется n подзадач, матрица $C = (c_{i,j})_{n \times n}$, элементы которой моделируют объемы данных, передаваемых между i -й и j -й подзадачами.

Вычислительную систему будем моделировать с помощью квадратной матрицы. Имеется n вычислительных узлов, матрица $D = (d_{i,j})_{n \times n}$, элементы которой моделируют стоимость обмена данными между i -м и j -м вычислительными узлами.

Решение будем записывать в виде перестановки p длины n , которая ставит в соответствие i -й подзадаче $p(i)$ вычислительный узел.

Критерий минимизирует суммарные издержки на пересылку данных:

$$\sum_{i=1}^n \sum_{j=1}^n c_{i,j} d_{p(i),p(j)} \rightarrow \min$$

В результате получили квадратичную задачу о назначениях [2], которая является NP-трудной [3].

Метод ветвей и границ

Этот метод является вариацией полного перебора с отсеком подмножеств допустимых решений, заведомо не содержащих оптимума.

Метод основан на обходе дерева, в каждой вершине которого зафиксирована определенная часть допустимого решения задачи. Посещая очередную вершину, алгоритм рассчитывает верхнюю и нижнюю оценки, в границах которых находится лучшее решение для данной ветви.

Верхняя оценка является достижимой, а нижняя – недостижимой. Способы нахождения этих оценок рассмотрены в [4].

Решения квадратичной задачи о назначениях записываются в виде перестановок.

Помимо точного метода ветвей и границ существует его модификации, дающие приближенное решение. Одна из модификаций, а также точный метод описан в [4].

В данной работе рассматривается улучшенная версия этой модификации. Ее суть заключается в следующем: начинается последовательный обход графа, до высоты обычно не более 1-3. После достижения этой высоты строятся соответствующие начальные перестановки, которые будут являться стартовыми для индивидуальных подзадач, которые исполняются параллельно. Каждая индивидуальная подзадача выполняет последовательный обход графа до фиксированной высоты. При ее достижении алгоритм останавливается, текущая верхняя оценка считается лучшим решением по направлению.

Все исполняющиеся параллельно подзадачи обновляют общее значение наименьшей верхней границы, которая используется для оценки перспективности обхода направления, а также по завершении своей работы записывают свое лучшее решение в общее множество решений.

По окончании работы всех подзадач выполняется поиск наилучшего решения из всего множества найденных решений.

Программная реализация и вычислительный эксперимент

Алгоритм реализован средствами языка C++.

Для хранения задачи используется класс `qap`. В нем хранятся: матрица данных параллельной программы, матрица стоимостей обмена данными между узлами.

Решение будем хранить в классе `permutation` – перестановка, индекс ячейки которой соответствует номеру задачи, а значение в этой ячейке – номер назначаемого вычислительного узла.

Критерий вычисляется в классе `calculator`. Этот класс хранит в себе указатели на соответствующие матрицы.

Функция `std::size_t calculator::criterion(permutation&permutation)` возвращает значение критерия для текущей перестановки.

Алгоритм запускается с помощью функции класса `qap`:

`solutionqap::get_solution(executorexec, lower_boundslower, upper_boundsupper, boolis_approximate)`, которая возвращает найденное решение задачи – перестановку и соответствующий ей критерий;

`enumexecutors` – перечисление, каждый элемент которого определяет используемый при решении «исполнитель»;

`enumlower_bounds` – перечисление, каждый элемент определяет способ вычисления нижней оценки;

`enumupper_bounds` – перечисление, каждый элемент определяет способ вычисления верхней оценки.

Получены следующие результаты:

Таблица 1

Результаты вычислений

№	Пример	Время работы, мс	Отклонение от оптимума
1	chr12a.dat	27	0 %
2	chr15a.dat	7477	0,56 %
3	chr15b.dat	1207	2,67 %
4	had14.dat	9746	0,72 %
5	had20.dat	1927	2,01 %
6	nug25.dat	1928	9 %
7	nug20.dat	1698	7,7 %

Эксперимент проводился на следующей машине: IntelCore i7 4700MQ @ 2.40 GHz, 8 GB, Windows 8.1.

Тестовые примеры взяты из публичной библиотеки QAPLIB [5].

Приближенная версия алгоритма позволила найти близкие к оптимуму решения, причем за небольшое время и больших размерностей, чем позволяла точная в [4].

Заключение

В работе рассмотрена проблема отображения параллельной программы на вычислительную систему. Проблема сформулирована в виде квадратичной задачи о назначениях. Рассмотрен и реализован параллельный приближенный метод ветвей и границ для поиска решения. Текущая реализация находит верхнюю оценку случайным образом и является недостаточно качественной. Поэтому в дальнейшем планируется замена алгоритма подсчета верхней оценки на генетический алгоритм.

Библиографический список

1. Старостин Н.В., Быкова М.А., Метод ветвей и границ решения квадратичной задачи о назначениях с приложениями с области высокопроизводительных вычислений, Системы управления и информационные технологии. 2017. № 1 (67). С. 13-18.
2. Koopmans T. C. and Beckmann M. J., Assignment problems and the location of economic activities, *Econometrica* 25, 1957, 53–76.
3. Батищев, Д.И. Вычислительная сложность экстремальных задач переборного типа: Учеб. пособие [Текст] /Д.И. Батищев, Д.И. Коган. – Нижний Новгород Изд-во Нижегород. ун-та: Нижегород. гос. ун-т им. Н. И. Лобачевского, 1994. – 114с.
4. Быкова М.А., Желтов С.А., Филимонов А.В., Решение квадратичной задачи о назначениях методом ветвей и границ, Математическое моделирование и суперкомпьютерные технологии. Труды XX Международной конференции (Н. Новгород, 23–27 ноября 2020 г.) / Под ред. проф. В.П. Гергеля. – Нижний Новгород: Изд-во Нижегородского госуниверситета, 2020, с.85-86.
5. Burkard R.E. QAPLIB – A Quadratic Assignment Problem Library [Text] /R.E. Burkard, S.E. Karisch, F. Rendl. – *Journal of Global Optimization*, Volume 10, Issue 4, 1997. – pp. 391–403

M.A. Bykova, S.A. Zheltov

APPROXIMATE ALGORITHM FOR SOLVING THE QUADRATIC ASSIGNMENT PROBLEM

Lobachevsky State University of Nizhny Novgorod

The paper considers the problem of mapping a parallel program to a computer system in order to optimize the execution time by reducing the time of data transfer. The problem is formulated as a quadratic assignment problem. A parallel approximate modification of the branch and bound method is developed and implemented.

Keywords: quadratic assignment problem, branch and bound method.

ЭВРИСТИЧЕСКИЙ АЛГОРИТМ РЕШЕНИЯ КВАДРАТИЧНОЙ ЗАДАЧИ О НАЗНАЧЕНИЯХ

Нижегородский государственный университет им. Н. И. Лобачевского

В работе был разработан и реализован эволюционно-генетический алгоритм для квадратичной задачи о назначениях. Также был разработан и реализован улучшающий алгоритм через поиск локальных минимумов. Были найдены близкие к оптимальным для известных тестовых примеров.

Ключевые слова: эволюционно-генетический алгоритм, квадратичная задача о назначениях, улучшающий алгоритм, локальный минимум.

Введение

В 1961 году Леон Штайнберг поставил задачу “проводки задней панели” [1]. Она состоит в том, чтобы расположить различные устройства на панели с предварительно подготовленными позициями, которые имеют некоторое расстояние между собой. Устройства на этих позициях соединяются друг с другом некоторым количеством проводов. Необходимо найти такое расположение устройств, чтобы суммарная длина проводов была минимальной. В годы развития компьютерной промышленности данная задача была одной из самых обсуждаемых в научном обществе. Несмотря на то, что проблема была поставлена около полувека назад, она не теряет своей актуальности – компьютерная промышленность еще развивается, а оборудование для вычислительных систем постоянно обновляется.

Данную задачу можно свести к квадратичной задаче о назначениях. Задача является NP-трудной [2], поэтому имеет смысл использовать эвристические алгоритмы. В данной работе был рассмотрен эволюционно-генетический алгоритм с последующими улучшениями решений.

Математическая постановка задачи

Даны 2 множества одинакового размера n и матрицы размера $n \times n$: $W = (w_{ij})$ и $D = (d_{ki})$, моделирующие связи между элементами каждого множества. В случае задачи “проводки задней панели” – это провода и расстояние между позициями. Необходимо найти взаимно-однозначное соответствие элементов заданных множеств, заданное перестановкой $p \in \Pi_n$, что целевая функция будет принимать минимальное значение. $\min_{p \in \Pi_n} \sum_{i=1}^n \sum_{j=1}^n w_{ij} d_{p(i)p(j)}$.

Эволюционно-генетический алгоритм

Эволюционно-генетический алгоритм представляет собой популяционно-генетический подход к решению задачи поиска, основанный на математическом моделировании биологических механизмов и процессов в живой природе с помощью принципов популяционной генетики. Эволюционно – генетический алгоритм не гарантирует обнаружения глобального оптимума за полиномиальное время, однако является наиболее простым в реализации среди прочих эвристических алгоритмов.

Выделяют 5 главных этапов алгоритма: формирование начальной популяции (множество начальных решений), воспроизводство новых особей через кроссовер, оценивание и обработка ограничений, формирование следующего поколения путем отбора [3].

В данной работе для кодирования решения была рассмотрено бинарное кодирование, которое представляет собой бинарную строку, кодирующую переход от текущего решения к новому [4, 8]. Таким образом на каждом поколении алгоритма исследуется разная окрестность текущего решения, и если найдется решение лучше текущего, то оно заменяется найденным. Также были рассмотрены классические операторы алгоритма, такие как: перестановочный кроссовер [5], β -турнир, утопия [5]. Кроме того, для данного алгоритма была реализована сортировка Шелла с последовательностью Седжвика [6], что значительно уменьшило время работы алгоритма, приведенного в [5]. Формирование начальной популяции и выбор родительских пар – случайные. Условие остановки алгоритма – лучшее решение не меняется на протяжении нескольких поколений. Результатом работы алгоритма будет вывод особи, с наилучшей приспособленностью, то есть такая особь, у которой значение целевой функции наименьше, среди всех рассмотренных в алгоритме. Разработанный алгоритм был программно реализован посредством C++.

Улучшение решения поиском локальных минимумов

После нескольких запусков эволюционного алгоритма с различными комбинациями значений β -коэффициента, количества родительских пар и интервалом перекрытия поколения возможно улучшение уже полученных решений. Полученные решение сортируются от лучших к худшим. Затем выбираются k первых полученных решений, и k начальных решений, соответствующих полученным.

После этого от каждого из начальных решений прокладывается “путь” к полученным, через попарные обмены в перестановках данных решений. На протяжении этого “пути” записываются новые решение, и, если среди них в этой последовательности будет найден локальный минимум, необходимо также совершить попарный обмен в его перестановке, для дальнейшего его улучшения, то есть уменьшения значения целевой функции. Как только не станет возможным уменьшать значение целевой функции путем попарного обмена, текущее решение становится частью результат работы “улучшающего алгоритма”. Конечным же результатом будет являться решение с наименьшим значением целевой функции из множества всех улучшенных локальных минимумов, найденных на k “путях”.

Ниже приведет визуальный пример работы данного алгоритма. $P_{нач}$ – начальное решение, $P_{кон}$ – конечное, зеленый знак “-” означает уменьшение решения на протяжении пути, красный знак “+” – увеличение, черными кругами выделены локальные минимумы, синими – решение, улучшение которого, путем попарного обмена невозможно.

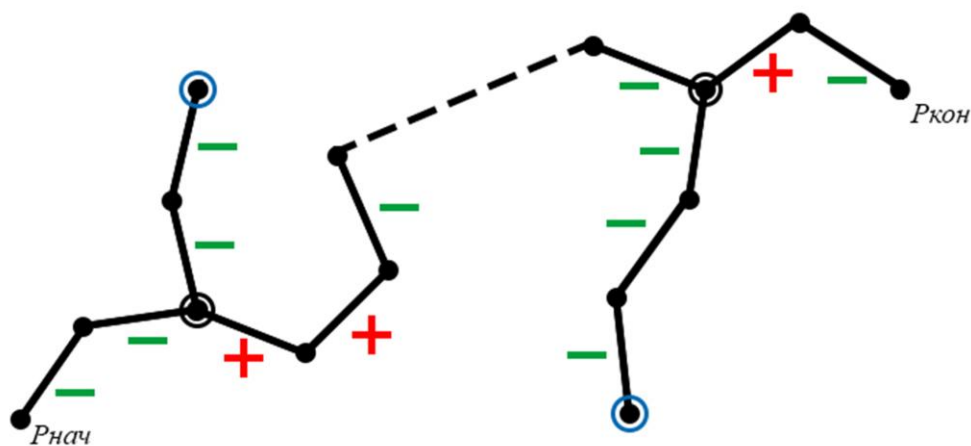


Рис. 1. Визуальный пример работы алгоритма

Вычислительный эксперимент

Для вычислительного эксперимента были выбраны известные задачи из QAPLIB [7]. Вычислительный эксперимент нацелен на оценку качества решений, получаемых через улучшение поиском локальных минимумов. Характеристики ПК, на котором производились эксперименты: ЦП – 8-ядерный процессор с тактовой частотой 4 ГГц (AMDFX-8350), ОЗУ – 14 Гб.

В таблице 1 приведены результаты работы эволюционно-генетического алгоритма на примерах размеров 25, 42, 60, 100 и разными конфигурациями алгоритма.

Таблица 1

Результаты эволюционно-генетического алгоритма

Размер	Найденное значение критерия	Отрыв от известного решения	Время (мс)
25	3794	1.32%	874
42	16400	3.59%	1647
60	7617460	5.40%	1112
100	22466100	6.33%	2503

Из полученных данных следует, что были найдены решения близкие к рекордному. Далее используем алгоритм улучшения.

Таблица 2

Результаты улучшающего алгоритма

Размер	Найденное значение критерия	Отрыв от известного решения	Время (мс)
25	3748	0.11%	64
42	15936	0.78%	386
60	7441240	3.24%	938
100	21560500	2.39%	8482

Улучшающий алгоритм путем нахождения локальных минимумов позволил найти близкие к оптимальным решения, причем решения, полученные данным алгоритмом более точные, чем в эволюционно-генетическом.

Заключение

В работе была рассмотрена и изучена задача “проводки задней панели”. Проблема была сформулирована в виде квадратичной задачи о назначениях. Разработаны и реализованы эволюционно-генетический алгоритм, а также улучшающий алгоритм через поиск локальных минимумов. Последний выдал решения, близкие к известным оптимальным. Средний отрыв составил 1.63%.

Библиографический список

1. Brixius N. W., Anstreicher K. M. The Steinberg Wiring Problem, 2004
1. Sahni S., Gonzalez T., P-complete approximation problems // Journal of the Association of Computing Machinery 23, 1976, P. 555-565
2. Батищев Д. И., Костюков В. Е., Неймарк Е. А., Старостин Н. В., Решение дискретных задач с помощью эволюционно-генетических алгоритмов. Нижний Новгород, 2009
3. Быкова М.А. Многоуровневые методы архитектурно-зависимой декомпозиции в области высокопроизводительных вычислений. Нижний Новгород, 2017
4. Быкова М.А., Хлопцев Н.А., Небайкин С.В., Решение квадратичной задачи о назначениях с помощью эволюционно-генетического алгоритма, Математическое моделирование и суперкомпьютерные технологии. Труды XX Международной конференции (Н. Новгород, 23–27 ноября 2020 г.) / Под ред. проф. В.П. Гергеля. – Нижний Новгород: Изд-во Нижегородского госуниверситета, 2020, с.87-88.

5. Sedgewick Robert, Algorithms in C. 1 (3rd ed.), Addison-Wesley, 1998, 3. 273–281
6. Burkard R.E., Karisch S.E., Rendl. F. QAPLIB – A Quadratic Assignment Problem Library // Journal of Global Optimization, Volume 10, Issue 4, 1997. – P. 391–403
7. Панкратова М.А., Гибридные схемы решения задачи отображения параллельной программы на вычислительную сеть // Системы управления и информационные технологии 3(61), Воронежский государственный университет, 2015, с. 64-70

М.А.Быкова, N.A. Khloptsev

HEURISTIC ALGORITHM FOR SOLVING THE QUADRATIC ASSIGNMENT PROBLEM

N. I. Lobachevsky State University of Nizhny Novgorod

Purpose: The article is devoted to research the effectiveness of the evolutionary-genetic algorithm and further improvement algorithm based on the search of local minimums.

Design/methodology/approach: The approach consists in application of the evolutionary-genetic algorithm on a set of problems with different sizes, testing the improvement algorithm, which takes in solutions from the evolutionary-genetic algorithm, and finally comparing the results of both algorithms.

Findings: The improvement algorithm showed good results, which lead to solutions close to known optimums. In problems with bigger size, the execution time of the improvement algorithm took even longer than in case of the evolutionary-genetic algorithm.

Research limitations/implications: The research indicates that the improvement algorithm requires further modifications and studying.

Originality/value: The combination of the evolutionary-genetic algorithm and the improvement algorithm have shown to be more useful, than the genetic algorithm alone. This combination may be used in finding solutions close to optimum in case of quadratic assignment problems.

Keywords: evolutionary-genetic algorithm, quadratic assignment problem, improvement algorithms, local minimums.

**В.В. Балашов, М.М. Годовицын, Ю.А. Живчикова, Н.В. Старостин,
А.В. Филимонов**

КОНЦЕПЦИЯ РЕАЛИЗАЦИИ СИСТЕМЫ ВЕРИФИКАЦИИ КОНСТРУКТИВНО-ТЕХНОЛОГИЧЕСКИХ ОГРАНИЧЕНИЙ ТОПОЛОГИИ МИКРОСХЕМ

Нижегородский государственный университет им. Н.И. Лобачевского

Изготовление микросхем – многоэтапный дорогостоящий процесс. Любое нарушение правил проектирования может повлечь за собой ошибки при изготовлении топологии (между фрагментами топологии могут появиться незапланированные соединения, некоторые фрагменты топологии могут изготовиться неправильной конфигурации и пр.). Для нахождения и устранения ошибок перед передачей топологии на изготовление необходимо выполнить её верификацию. Первым этапом верификации является проверка на соответствие нормам конструктивно-технологических ограничений, гарантирующая, что все фрагменты топологии с учетом особенностей оборудования конкретной фабрики могут быть изготовлены и впоследствии будут работать корректно. В статье рассматриваются требования к системе верификации конструктивно-технологических ограничений топологии микросхем, подходы к реализации и архитектура системы.

Ключевые слова: верификация, топология микросхем, конструктивно-технологические ограничения

Введение

При производстве микросхем с субмикронными проектными нормами сроки и стоимость изготовления крайне важны, поэтому задача получения работоспособного кристалла перед передачей спроектированной топологии на производство имеет большое значение. Необходимо проверить топологию на соответствие нормам конструктивно-технологических ограничений (КТО), обеспечивающим возможность качественного изготовления. Эту задачу решают с помощью САПР для проверки норм КТО (DRC). КТО - разрешенное взаимное расположение топологических элементов и топологических структур с учетом конструктивных особенностей и возможностей технологического процесса. Взаимное расположение элементов топологии приводится в документе «Правила проектирования», который содержит описание совокупности норм, ограничений, правил и процедур, представляемых в установленных форматах, разрабатываемых для заданного технологического процесса, соблюдение которых при проектировании кристалла интегральной микросхемы обеспечивает возможность его изготовления.

На основании информации, содержащейся в документе, разрабатывается файл правил для САПР, содержащий правила проверки норм КТО. Правило проверки — набор операций, объединенных именем, имеющих описание и выдающих результат. По результатам проверки разработчиком принимается решение о доработке топологии микросхемы или передаче её на изготовление.

Лидирующие позиции в сегменте верификации топологии интегральных схем на рынке EDA (Electronic Design Automation) делят между собой следующие платформы [1]:

- Hercules и IC Validator ф. Synopsys (США);
- Diva, DRACULA, Assura, PVS и Pegasus ф. Cadence Design Systems (США);
- Calibre ф. Siemens DISoftware (Германия).

Фактически, разработка и развитие ведущих коммерческих продуктов в данной сфере обеспечивается зарубежными компаниями. В соответствии со стратегией импортозамещения в ИТ-отрасли РФ требуется разработка отечественных программных продуктов в области дизайна и верификации топологии микросхем.

Требования к системе верификации конструктивно-технологических ограничений топологии микросхем

Создание системы верификации КТО предполагает разработку общей структуры и информационной модели функционирования базовых компонентов системы верификации, КТО топологии микросхем, обеспечивающих чтение данных топологии ИС и правил проверки норм КТО, генерацию БД с результатами проверки норм. Очевидно, что в рамках данной активности должны быть выбраны интерпретаторы языка входного файла правил и программные интерфейсы, обеспечивающие взаимодействие с внешними подсистемами (такими как программа визуализации результатов проверки КТО, редактор топологии и пр.). Необходимо реализовать маршрут проведения физической верификации топологии. При этом система верификации должна обеспечивать следующие возможности:

- чтение исходной информации (слоев) из файла GDSII или из базы данных ранее сохраненных слоев;
- описание процесса проверки топологии с использованием ряда общеупотребимых функций работы над слоями, таких как:
 - выбор примитивов (прямоугольники, ребра, многоугольники), удовлетворяющих определенным условиям;
 - генерация новых слоев путем совершения операций над существующими (объединение, пересечение и т.д.);
- контроль описания процесса проверки с точки зрения синтаксической и логической корректности;
- сохранение результатов верификации в специальную БД, включая отчеты работы алгоритмов верификации для их дальнейшей обработки/представления пользователю.

При создании архитектуры системы верификации КТО топологии микросхем возникает важный концептуальный вопрос о выборе подхода к разработке: воспроизводить концепции существующих программных продуктов, либо искать новые способы решения поставленных задач.

Подходы к проверке конструктивно-технологических ограничений

В настоящий момент основополагающим элементом верификации является файл правил. Используемые промышленные форматы представления файлов правил реализуют декларативный подход к описанию правил проверки конструктивно-технологических ограничений (например, коммерческий формат SVRF ф. Siemens DISoftware [2]). Т.е. файл правил содержит описание норм в виде перечисления правил (пример: расстояние между прямоугольниками в слое не должно превышать заданной величины), что предполагает последующую разработку и реализацию компонентов для распознавания описанных команд, преобразования их в программную модель, выполнения описанных проверок и т.д.

Ввиду большого количества возможных проверок, увеличивающегося с возрастанием сложности проектируемых интегральных схем [3], применение декларативного подхода к описанию файлов проверки правил является нецелесообразным при разработке системы верификации КТО в современных производственных реалиях. Вместо этого предлагается императивный путь: создание набора инструкций, описывающих процесс верификации КТО топологии микросхемы от загрузки данных до выгрузки результатов верификации (в частности, для правила проверки: выбери в слое прямоугольники, измерь расстояние между ними – эта величина не должна превышать

заданное значение). Результирующий скрипт будет содержать готовый алгоритм проверки (скрипт верификации). Данный подход в дальнейшем требует лишь интерпретации инструкций скрипта.

Для реализации императивного подхода предлагается использовать скриптовый язык Lua как средство описания алгоритма проверки. Процедурный динамически типизированный модульный язык Lua обладает простым дизайном, широким набором доступных библиотек и эффективными средствами межъязыкового взаимодействия, ориентированными, преимущественно, на вызов библиотек языка C и на работу в C-окружении [4], что позволяет вести разработку программных модулей с его использованием в рамках разработанной архитектуры системы верификации конструктивно-технологических ограничений топологии микросхем.

Архитектура системы верификации конструктивно-технологических ограничений топологии микросхем

Предлагаемая к реализации структура системы верификации конструктивно-технологических ограничений топологии микросхем приведена на рис. 1.

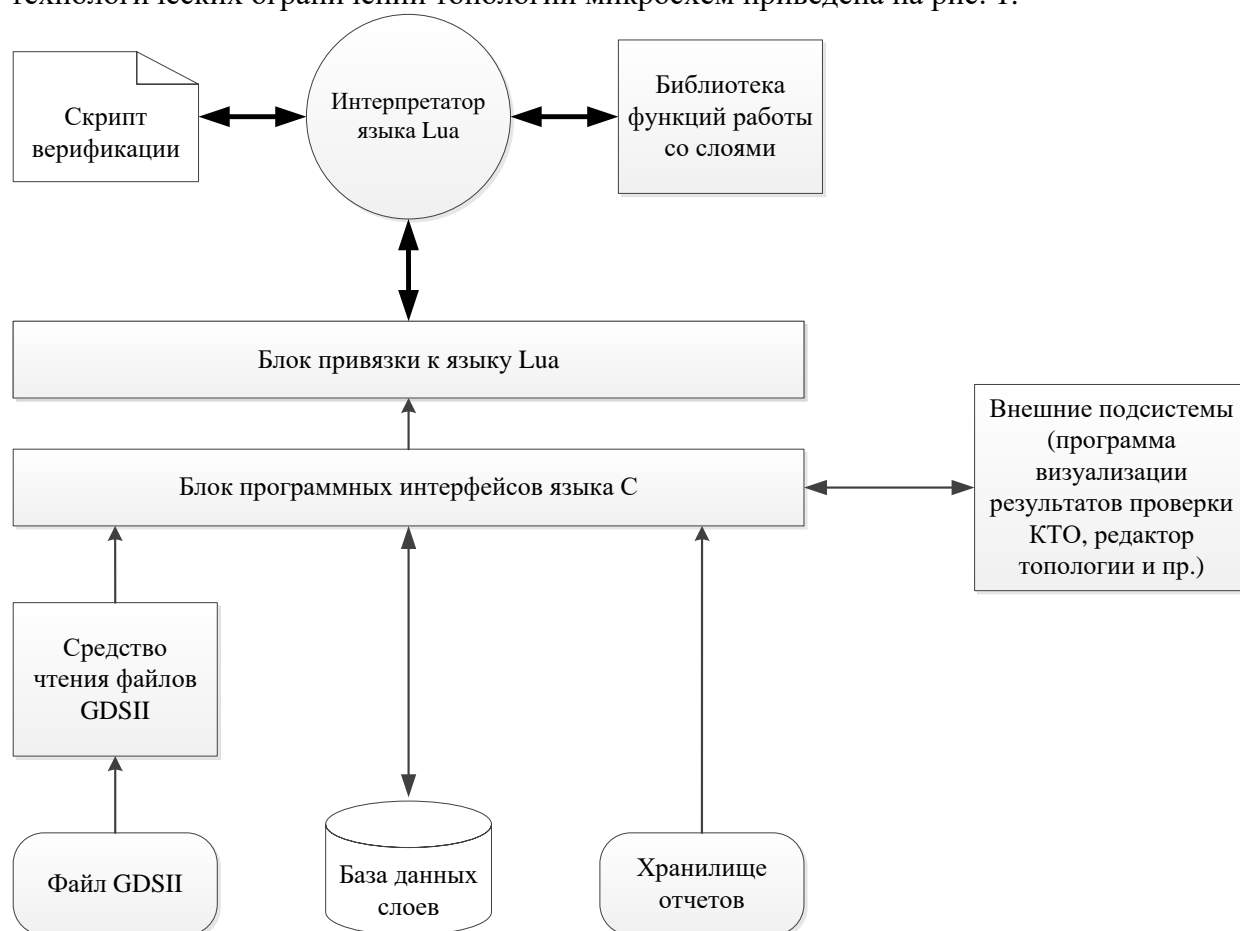


Рис. 1. Архитектура системы верификации конструктивно-технологических ограничений топологии микросхем

В соответствии с предлагаемой архитектурой система верификации конструктивно-технологических ограничений топологии микросхем включает следующие компоненты:

- интерпретатор языка Lua;
- библиотека функций для работы со слоями;
- блок работы с файлами GDSII в составе:
 - a. средство чтения файлов GDSII;
 - b. блок программных интерфейсов языка C;
 - c. блок привязки к языку Lua;

- блок работы с БД слоев в составе:
 - a. база данных слоев;
 - b. блок программных интерфейсов языка C;
 - c. блок привязки к языку Lua;
- блок работы с отчетами в составе:
 - a. хранилище отчетов работы алгоритмов верификации (логов);
 - b. блок программных интерфейсов языка C.

Следует отметить, что входными данными системы верификации КТО являются GDSII файл топологии интегральной схемы и скрипт верификации. Выходными данными являются БД результирующих слоев и отчеты работы алгоритмов верификации.

Таким образом, для скрипта верификации обеспечиваются следующие возможности:

- чтение слоев из файла GDSII;
- обработка слоев при помощи функций из библиотеки функций для работы со слоями;
- сохранение слоев в базе данных слоев;
- сохранение логов и отладочной информации в хранилище отчетов работы алгоритмов верификации.

Предлагаемая архитектура системы предоставляет широкие возможности с точки зрения расширяемости. Другими достоинствами являются простота и большая описательная сила языка Lua в рамках императивного подхода к разработке системы.

Заключение

В контексте реализации отечественной системы верификации КТО предложен императивный подход к верификации, КТО, отличающийся идеологической новизной. На основании сформулированных требований к системе верификации конструктивно-технологических ограничений топологии микросхем разработана архитектура системы, предусматривающая компоненты для чтения входных данных и работы со слоями топологии, интерпретации входного скрипта проверки, работы с БД слоев топологии, работы с отчетами верификации.

Библиографический список

1. Павлов С.И. Системы высокопроизводительных вычислений в 2019–2020 годах: обзор достижений и анализ рынков Часть V. Сфера PLM, включая CAE и EDA / С.И. Павлов // CAD/CAM/CAE Observer. – 2020. – №7 (139). – С. 4-19.
2. Calibre Verification User's Manual. – URL: <http://en.pudn.com/Download/item/id/1028073.html> (дата обращения: 28.12.2020).
3. Введение в системы автоматизированного проектирования интегральных схем: учебно-методическое пособие. Ч. 1 / сост.: А.В. Тучин, Е.Н. Бормонтов, К.Г. Пономарев. – Воронеж: Издательский дом ВГУ, 2017. – 111 с.
4. Lua 5.4 Reference Manual. – URL: <https://www.lua.org/manual/5.4/> (дата обращения: 28.12.2020). – Текст: электронный.

CONCEPT OF THE IC LAYOUT DESIGN RULE CHECKING SYSTEM IMPLEMENTATION

Nizhniy Novgorod state university n.a. N.I. Lobachevskiy

Purpose: This article considers requirements for the IC layout design rule checking (DRC) system, existing and possible approaches to the implementation and also proposes architecture of the DRC system.

Design/methodology/approach: Imperative approach is proposed for DRC system architecture developing instead of existing declarative method. Declarative method implies creating a rule file, containing norms description as a listing of design rules. Imperative approach implies creating a script, containing set of instructions, which describes verification process. This approach later requires interpretation script instructions only.

Findings: New imperative approach to the DRC system designing is proposed. Architecture based on this approach includes components for the input data reading, working with the IC layout layers, interpretation of the verification script, working with the layer database and logs.

Research limitations/implications: The present article provides a starting-point for further research in the DRC system developing.

Originality/value: DRC system based on propose darchitecture can be applied for layout verification atIC designing enterprises. Moreover, there aren't such systems in the local IT market. Therefore, DRC system is really marketable product today.

Keywords: verification, IClayout, design rule checking

УЧЕТ НЕОДНОРОДНОЙ ВМЕСТИМОСТИ ПОДОБЛАСТЕЙ В СИЛОВЫХ СХЕМАХ РАЗМЕЩЕНИЯ ГРАФА

*Национальный исследовательский Нижегородский государственный
университет им. Н.И. Лобачевского*

Статья посвящена рассмотрению задачи укладки графа на плоскости с учетом наличия подобластей ограниченной вместимости. Производится формальная постановка задачи с учетом данных ограничений. Поиск решения задачи предлагается построить по гибридной схеме, основанной на применении известного подхода силовой укладки графа в сочетании с перераспределением вершин при помощи итерационной трансформации пространства размещения. Рассмотренная методика решения задачи реализована программно, приведены результаты вычислительных экспериментов.

Ключевые слова: силовая укладка графа, учет вместимости подобластей, многоуровневая схема.

Введение

В различных областях науки и прикладных задачах возникает проблематика, связанная с размещением (укладкой) графов на плоскости. В частности, подобные задачи возникают на этапах физического проектирования интегральных схем [1,2]. Так, например, в задаче глобального размещения требуется указать размещение компонент интегральной схемы с точностью до областей монтажного пространства кристалла, где каждая область представлена компактной группой посадочных мест. В данной задаче важно не только минимизировать общую оценку длины трасс, но и обеспечить потенциальную возможность последующего бесконфликтного назначения компонентов на посадочные места. Приведем формальную постановку данной задачи.

Постановка задачи

В качестве исходных данных выступает взвешенный граф $G(V, E, w, u)$ схемы, где V – множество компонент схемы, $V = \{v_1, \dots, v_n\}$; E – множество связей схемы (ребро между парой вершин означает наличие цепи, связывающих контакты соответствующих компонент), $E \subseteq V^{(2)}$, $|E| = m$; $w = (w_1, \dots, w_n)$, где w_i – число занимаемых посадочных мест i -м компонентом, $w_i \in N, i = \overline{1, n}$; $u = (u_1, \dots, u_m)$, где u_j – число общих цепей, в которые входят инцидентные компоненты j -го ребра, $u_j \in N, j = \overline{1, m}$.

Монтажное пространство кристалла моделируется набором непересекающихся прямоугольных областей $(R_l, \rho_l), l = \overline{1, k}$. Здесь R_l соответствует некоторой прямоугольной области $A_l = [X_l^-, X_l^+] \times [Y_l^-, Y_l^+] \subset R^2$, а значения ρ_l указывают на число доступных посадочных мест в данной области, $\rho_l \in N$.

В задаче требуется для каждого компонента v_1, \dots, v_n указать размещение (p_1, \dots, p_n) в рамках монтажного пространства кристалла, где $p_i = (x_i, y_i) \in R^2, i = \overline{1, n}$.

Потребуем, чтобы вершины могли размещаться строго в рамках заданных прямоугольных областей

$$p_i \in A, i = \overline{1, n}, \text{ где } A = \bigcup_{l=1}^k A_l. \quad (1)$$

Учтем потенциальную возможность последующего бесконфликтного назначения компонентов на посадочные места, для чего ограничим суммарное число посадочных мест, которые занимают компоненты, размещенные в прямоугольные области

$$\sum_{v_i: p_i \in A_l} u(v_i) \leq \rho_l, l = \overline{1, k}. \quad (2)$$

Для оценки качества размещений p выберем следующие обобщённые функции цели.

Во-первых, важно в результате размещения получить компактные трассы цепей, для чего будем использовать правило «чем сильнее связаны компоненты, тем ближе они должны располагаться друг к другу», реализованное в виде функционала

$$F_a(p) = \sum_{(i,j) \in E} w_{ij} \|p_i - p_j\|^a \rightarrow \min, \text{ где } a \geq 0. \quad (3)$$

Во-вторых, дополнительно учтём возможность последующего бесконфликтного размещения компонентов, для чего воспользуемся следующим правилом - «чем больше требуется посадочных мест компонентам, тем на большем расстоянии друг от друга они должны находиться», реализованное в виде функционала

$$F_r(p) = \sum_{(i,j) : i \neq j} u_i u_j \|p_i - p_j\|^b \rightarrow \min, \text{ где } b \leq 0. \quad (4)$$

В общем случае функции целей (3) и (4) противоречивы. В случае многокритериальной постановки решением является не единственное размещение, а множество решений, оптимальных по Парето. Предлагается рассмотреть данную задачу в однокритериальной постановке, полученной классической линейной сверткой:

$$\lambda F_a(p) + (1 - \lambda) F_r(p) \rightarrow \min, \text{ где } \lambda \in [0,1]. \quad (5)$$

Задачу (1)-(5) будем называть обобщенной задачей размещения графовых структур на плоскости. В представленных функциях цели (3),(4) параметры a и b позволяют управлять «значимостью» правил. В данной работе были использованы следующие значения параметров: $a = 1, b = -2$.

Метод решения

Задача (5) относится к классу задач раскроя и упаковки (cutting and packing problem) и является NP-трудной задачей [3], поэтому получение точного решения за полиномиальное время в общем случае не представляется возможным.

Разобьем решение задачи на два этапа:

1. Получение размещения вершин графа без учета ограничений (1)-(2) с помощью алгоритмов, основанных на силовой укладке графа [4, 5].
2. Учет ограничений (1)-(2) при помощи итерационной трансформации пространства размещения.

На первом этапе для поиска решения задачи используется многоуровневая схема [6] в сочетании с алгоритмом силовой укладки графа [4,5]. Центральной идеей многоуровневого метода является последовательное редуцирование размерности задачи путем обобщения информации о задаче и включения в редуцированный вариант наиболее существенных данных, решение редуцированной задачи с помощью алгоритма силовой укладки и последовательное восстановление полученного решения до размерности исходной задачи. При этом на каждом этапе восстановления выполняется улучшение восстанавливаемого решения с учетом добавления ранее исключенных из рассмотрения данных. Для улучшения решения также используется силовой алгоритм. Многоуровневые методики широко применяются при решении различных задач на графах, например, задач декомпозиции графов [7,8]. Силовая укладка графа выполняется с помощью алгоритма, описанного в [5]. Результатом первого этапа решения задачи является размещение вершин графа на плоскости, удовлетворяющее функционалу (5).

На втором этапе для учета ограничений (1)-(2) используется алгоритм, в основу которого положена концепция последовательной трансформации пространства размещения. Суть идеи в построении отображения $f: R^2 \rightarrow R^2$, которое трансформирует пространство размещения таким образом, что области размещения, где сконцентрировано наибольшее количество вершин графа, расширяются за счет регионов с меньшим количеством.

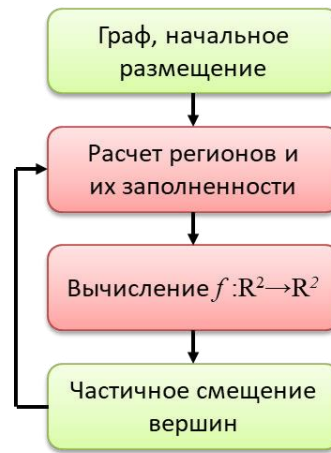


Рис. 1. Схема алгоритма трансформации пространства размещения.

Непосредственно алгоритм трансформации пространства построен по итерационной схеме и состоит из нескольких шагов (рис. 1). Сначала область размещения графа разбивается равномерной решеткой размера $r \times r$ на r^2 регионов $R_{s,t} \subset R^2$, $s, t = \overline{1, r}$, здесь r – параметр алгоритма. Далее для каждого региона вычисляется оценка его вместимости, которая определяется через относительную вместимость областей A_i , пересекающихся с регионом

$$\rho(R_{s,t}) = \sum_{i=1}^k \rho_i \frac{SQUARE(R_{s,t} \cap A_i)}{SQUARE(A_i)}, \quad s, t = \overline{1, r}, \quad (6)$$

где $SQUARE()$ – функция, возвращающая площадь от области размещения.

На каждой итерации алгоритма для каждого региона будем вычислять заполненность региона – суммарный вес вершин, попавших в результате размещения в данный регион.

$$\sigma(R_{s,t}) = \sum_{\substack{p_i \in R_{s,t} \\ i=\overline{1, n}}} u_i, \quad s, t = \overline{1, r}. \quad (7)$$

Введем новую характеристику – вес региона, которая показывает отношение наполненности к оценке вместимости региона

$$\tau(R_{s,t}) = \frac{\sigma(R_{s,t})}{\rho(R_{s,t})}, \quad s, t = \overline{1, r}. \quad (8)$$

Значение характеристики $\tau > 1$ для некоторого региона указывает на избыточное количество вершин. Значение $\tau < 1$ сигнализирует о наличии резерва для размещения в области дополнительных вершин. Возникает задача построения такой трансформации исходной равномерной решетки, при которой регионы с избытком вершин будут увеличены за счет регионов с резервами по вместимости. Построение трансформированной сетки состоит в расчете такого положения узлов решетки, при которой веса регионов и размеры регионов приведены в соответствие – регионам с большим весом соответствует большие ячейки решетки по отношению к остальным регионам. На основе построенной сетки выполняется пересчет нового расположения вершин методом линейной интерполяции. Последовательно процесс трансформации пространства размещения изображен на рис. 2.

Данный процесс реализован в итерационной схеме – трансформация происходит не единовременно, а многократно до тех пор, пока не будут выполнены ограничения (1), (2).

Описанная концепция решения задачи была реализована программно с использованием платформы Microsoft .NET (C#) в виде автоматизированной системы, позволяющей пользователю контролировать процесс генерации решения в части последовательности выполнения этапов, прерывания процесса расчета и досрочного перехода к следующему или возврату к выполнению предыдущего этапа, а также повторного запуска любого из них. Такой подход, варьируя параметр $\lambda \in [0, 1]$, позволяет фактически построить решения из области Парето, удовлетворяющих, с одной стороны,

критериям по взаимному расположению вершин в зависимости от их весов и связей между ними, с другой стороны –ограничениям на вместимость подобластей в пространстве размещения.

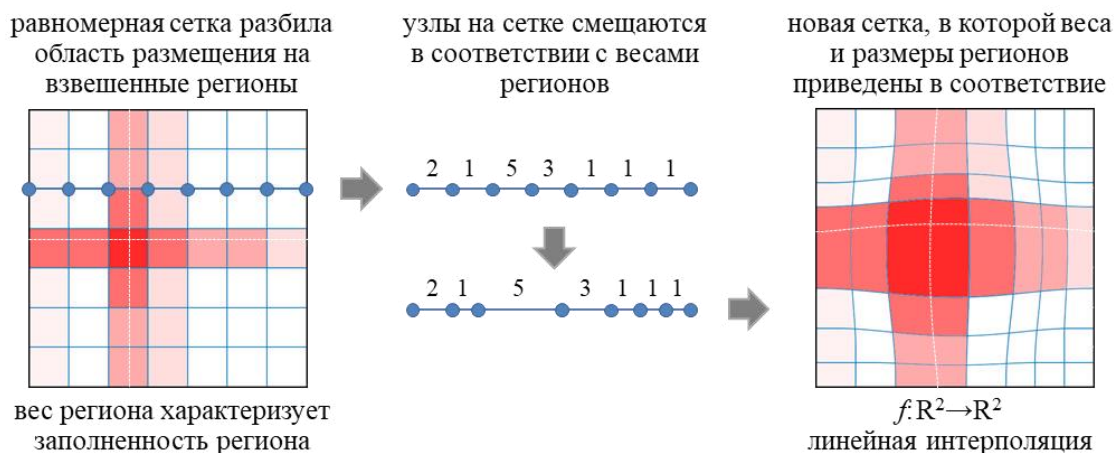


Рис. 2. Процесс трансформации пространства размещения

В качестве тестовой базы использованы регулярные графы размером от 10000 до 40000 вершин. На рис.3 приведены этапы работы программной системы при построении укладки сеточного графа (9947 вершин, 39336 ребер), при этом в пространстве размещения задана область запрета с нулевой вместимостью вершин (изображена в виде белого прямоугольника). На рисунке слева показан результат размещения графа без учета ограничений на вместимость подобластей. На рисунке справа – результат перераспределения вершин графа, в котором произведен учет неравномерной вместимости подобластей.

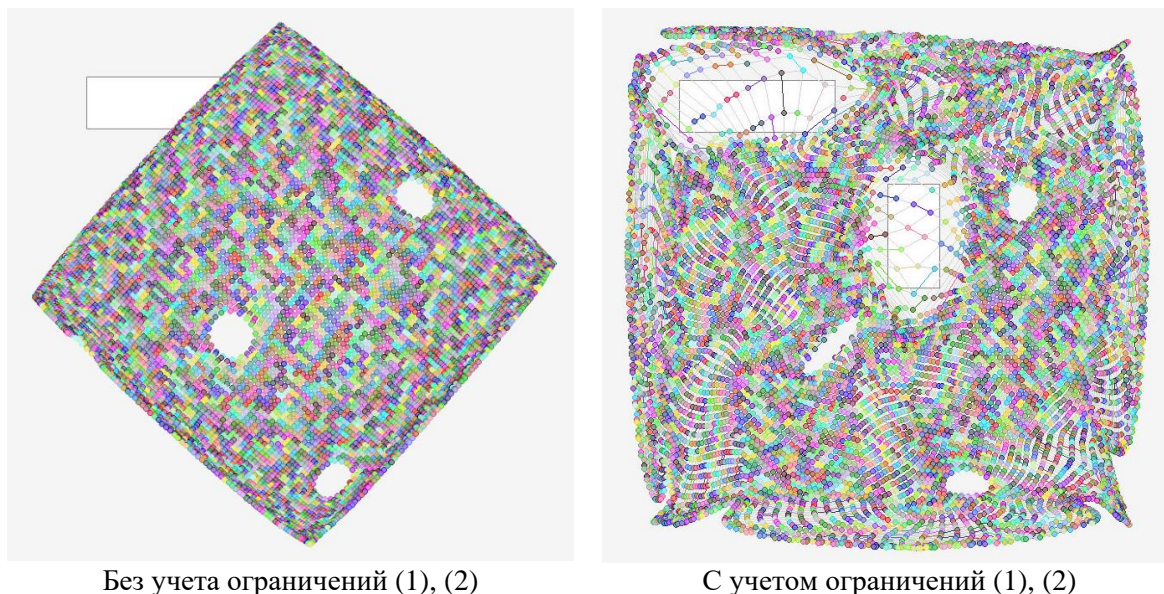


Рис. 3. Результаты размещения графа

Заключение

В рамках данной работы была исследована проблема поиска решений для класса задач размещения взвешенных графовых структур на областях с неравномерной вместимостью. Была предложена эвристическая двухэтапная итерационная схема решения задачи, основанная на классических алгоритмах силовой укладки графа и технологии трансформации пространства размещения. Экспериментально подтверждено, что данный

подход позволяет эффективно генерировать размещения большиеразмерных графов и может быть рекомендован в качестве метода решения задач глобального размещения компонент интегральных схем.

Библиографический список

1. Старостин Н.В., Филимонов А.В., Балашов В.В. Решение задачи размещения элементов, специализированных больших интегральных схем на основе базовых матричных кристаллов. Системы управления и информационные технологии. 2009. № 2-1 (36). С. 189-194.
2. Батищев Д.И., Старостин Н.В., Филимонов А.В. Многоуровневый алгоритм решения задачи компоновки интегральных схем. Системы управления и информационные технологии. 2007. Т. 29. № 3. С. 48-52.
3. Гэри М., Джонсон Д. Вычислительные машины и труднорешаемые задачи. М.: Мир, 1982. – 416 с.
4. Старостин Н.В., Небайкин С.В., Волков В.О., Басалин П.Д. Быстрые реализации алгоритма силовой укладки графа. В сборнике: Информационные системы и технологии ИСТ-2017. Материалы докладов XXIII Международной научно-технической конференции, посвященной 100-летию НГТУ. 2017. С. 758-761.
5. Старостин Н.В., Небайкин С.В. Аспекты использования алгоритма силовой укладки графа в промышленной технологии укладки кабелей в судостроении. Труды НГТУ им. Р.Е. Алексеева. 2019. № 2 (125). С. 70-78.
6. Старостин Н.В., Небайкин С.В., Волков В.О. Концепция многоуровневости как инструмент ускорения классического алгоритма силовой укладки графа. Информационные технологии моделирования и управления. 2018. Т. 109. № 1. С. 40-46.
7. Starostin N.V., Vykova M.A., Nebaikin S.V. Multilevel procedure for decomposition and mapping graphs. В сборнике: Journal of Physics: Conference Series. Krasnoyarsk Science and Technology City Hall of the Russian Union of Scientific and Engineering Associations. Krasnoyarsk, Russian Federation, 2020. С. 32017.
8. Батищев Д.И., Старостин Н.В., Филимонов А.В. Многоуровневый генетический алгоритм решения задачи декомпозиции гиперграфа. Известия СПбГЭТУ ЛЭТИ. 2007. № 1. С. 3-13.

S.V. Nebaikin

TREATMENT NON-UNIFORM CAPACITY OF SUB-REGIONS IN FORCE-DIRECTED GRAPH PLACEMENT

Lobachevsky State University of Nizhny Novgorod

The article considers the problem of creating a graph layout on a plane, taking into account the presence of sub-regions with limited capacity. A formal statement of the problem is carried out according to these restrictions. Solution search is proposed to construct as a hybrid scheme based on the well-known approach of force directed graph placement combined with vertices relocation employing placement space iterative transformation. The considered method of solving the problem is implemented in software automated system, the results of computational experiments are presented.

Keywords: force-directed graph layout, treatment non-uniform capacity of sub-regions, multilevel scheme.

**ПОСТРОЕНИЕ ЕДИНОЙ ТРИАНГУЛЯЦИИ
ПОВЕРХНОСТИ ГЕОМЕТРИЧЕСКОГО ТЕЛА**

Нижегородский государственный университет им. Н.И. Лобачевского

В предметной области инженерного анализа важной задачей является работа с сеточным представлением геометрических объектов, в частности твердых тел. Одной из актуальных проблем является получения сеточного представления поверхности геометрического тела при наличии независимых сеточных представлений (триангуляций) граней тела. В работе задача объединения сеточных представлений рассматривается для случая, когда сеточные представления являются триангуляциями. Представлен последовательный алгоритм объединения нескольких триангуляций в единую. Выделены и описаны основные этапы данного алгоритма. Произведен анализ работы алгоритма на практических моделях различной категории сложности.

Ключевые слова: сеточное представление, триангуляция, инженерный анализ.

Введение

Для моделирования прочности в инженерном анализе необходимо наличие сеточного представления объекта, в связи с чем возникает задача его построения [1, 2]. Также при работе с сеточным представлением, его редактированием и конвертацией в различных системах обработки возможно возникновение ошибок в структуре данных, которые могут привести к таким проблемам как: дубликаты точек, потеря информации об узлах и связях сеточного представления и другие, что в итоге приведет к некорректному сеточному представлению данного объекта и невозможности его использования в дальнейшем. В связи с этим является актуальной задача восстановления информации о сеточном представлении геометрического тела, состоящая в том, что при наличии триангуляций участков тела (граней), нужно получить единое сеточное представление всего тела.

Постановка задачи

Пусть задан конечный набор точек $\{P_i\}, i = \overline{1, n}$ в пространстве R^3 . Триангуляцией данного набора точек будет являться набор треугольников $\{T_j\}, j = \overline{1, m}$, который будет удовлетворять следующему набору условий:

- каждая точка P_i является вершиной хотя бы одного из треугольников T_j ;
- каждая точка P_i из всего набора точек должна являться уникальной, то есть не должно существовать ее дубликатов;
- каждый треугольник T_j содержит только три точки P_i из данного набора точек, а каждая из этих точек является вершиной данного треугольника;
- каждый треугольник T_j должен быть невырожденным (все три вершины треугольника не могут находиться на одной прямой).

Два треугольника называются соседними, если они имеют общее ребро. Каждый из треугольников может иметь не более трех соседей.

Граничным треугольником триангуляции будет называться такой треугольник T_j , у которого хотя бы с одной из его сторон отсутствует соседний треугольник, то есть среди всего набора точек триангуляции имеется такая пара точек P_i , которая одновременно может принадлежать только ему. Если у треугольника присутствуют все три соседа, то он называется внутренним. Пример граничного и внутреннего треугольника показан рис. 1 (граничный треугольник выделен черным, внутренний темно-серым).

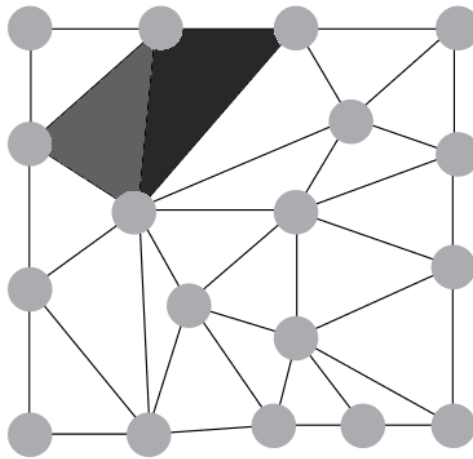


Рис. 1. Пример внутреннего и граничного треугольника триангуляции

Границей триангуляции Γ называется последовательный набор точек P_i , которые являются вершинами граничных треугольников T_j и расположены на сторонах с отсутствующими соседними треугольниками.

Гранью геометрического тела G называется двухпараметрическая поверхность $S(u, v)$ в пространстве R^3 , границы которой представляются в виде набора однопараметрических кривых $f_i(t)$.

Швом H грани G называется такая пара кривых $f_i(t), f_j(t)$, которые являются совпадающими в пространстве R^3 .

Триангуляцией \mathcal{T} грани G называется триангуляция набора точек, каждая из которых лежит в области грани G и с некоторой степенью точности аппроксимирует ее поверхность.

В качестве исходных параметров данной задачи выступают сеточные представления каждой грани геометрического тела и точность геометрической модели.

Ограничения данной задачи включают в себя следующие условия: триангуляции каждой грани геометрического тела не должна включать в себя дубликаты точек; не должно существовать пересечений каких-либо триангуляций по их внутренним треугольникам.

Результатом решения данной задачи является единая триангуляция, являющаяся результатом объединения всех исходных триангуляций и покрывающая всю поверхность геометрического тела. Две триангуляции объединяются друг с другом в единую по их ближайшей части границы таким образом, что результирующая триангуляция \mathcal{T} будет состоять из множества точек и треугольников обеих этих триангуляций, которые могут быть скорректированы в ходе их объединения, как это показано на Рис. 4 и Рис. 5. То есть в ходе сшивания обеих триангуляций может быть изменено как множество точек (например, добавлены новые, или удалены какие-либо из старых точек), так и множество треугольников (изменены вершины общих треугольников, добавлены новые треугольники и/или удалены старые) [3].

Алгоритм построения общей триангуляции

Алгоритм объединения нескольких триангуляций в единую представляет собой последовательную сшивку сеточных представлений всех граней тела до тех пор, пока общая триангуляция не будет покрывать его полностью. Это достигается при помощи циклического присоединения сеточных представлений граней тела к начальной триангуляции (пустой по умолчанию), т.е. задача сводится к объединению двух триангуляций. В процессе сшивания двух триангуляций можно выделить следующие основные этапы:

- выделение границ;
- поиск общей границы;
- корректировка общей границы;
- сшивка триангуляций.

Выделение границ триангуляций представляет собой отделение граничных треугольников от внутренних и дальнейший последовательный проход по сторонам граничных треугольников. Берется случайный граничный треугольник, в границу триангуляции добавляются его вершины со стороны без соседа. Далее происходит поиск треугольников, которые содержат последнюю добавленную точку. Так как она может входить сразу во множество треугольников, то необходимо отобрать среди всех именно тот, который является граничным и еще не был рассмотрен. Процесс продолжается до тех пор, пока последняя добавленная точка не совпадет с первой точкой границы и все граничные треугольники не будут рассмотрены. Если, например, был совершен полный обход границы триангуляции и остались не рассмотренные граничные треугольники, то это означает, что триангуляция имеет несколько границ (есть «отверстие» внутри нее), в таком случае запоминается текущая граница и начинается выделение новой. Пример выделения границы показан на рис. 2 (граничные точки отмечены темно-серым).

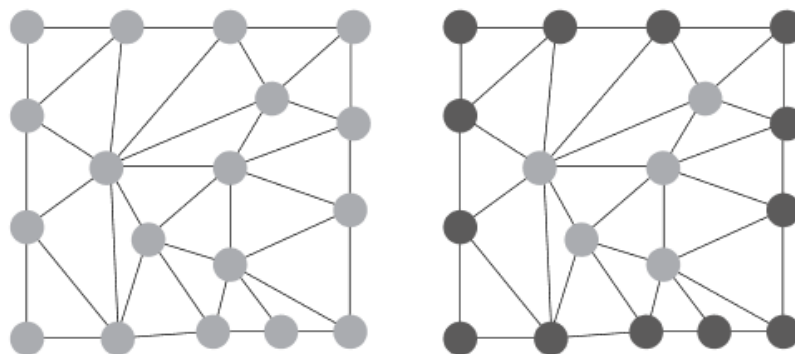


Рис. 2. Выделение границы триангуляции

Поиск общей границы. Для дальнейшего объединения двух триангуляций необходимо выяснить как они располагаются друг относительно друга. Например, две триангуляции могут быть близки сразу по всей длине своих границ, или только по ее части, а могут и вовсе быть расположены далеко друг от друга. Для этого необходим алгоритм поиска ближайших частей границ триангуляции. Сначала ищется пара ближайших сегментов обеих триангуляций, расположенных на границах. Далее осуществляется последовательный проход по выделенным границам, начиная с найденных ранее сегментов, и анализом их расположения друг относительно друга. Если два следующих сегмента границ после стартовых являются близкими, то их индексы запоминаются и рассматривается следующая пара сегментов. Если же во время последовательного прохода по сегментам попадает такая пара сегментов, которые являются далекими, то алгоритм переходит вновь на стартовые сегменты и совершает проход в противоположную сторону по границам. Это делается на случай, когда стартовые сегменты находятся в середине близких участков границ. Как только был совершен последовательный проход в обе стороны от стартовых сегментов вдоль границ двух триангуляций, то производится поиск следующей пары ближайших сегментов. Если же все сегменты границ триангуляций были рассмотрены, то работа алгоритма завершается. В результате, на границах триангуляций будут выделены участки, расположенные наиболее близко друг к другу (рис. 3).

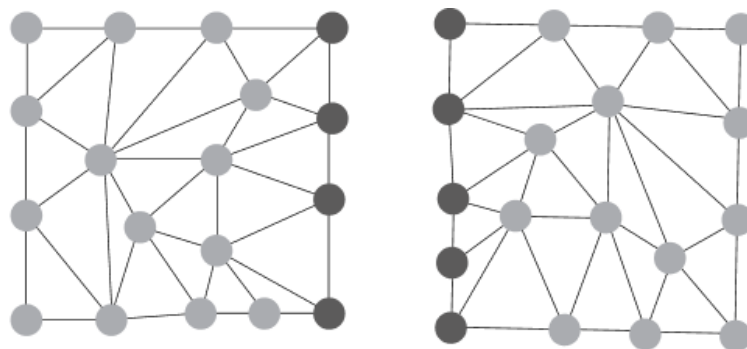


Рис. 3. Поиск общих частей границ триангуляций

Корректировка общей границы. Как только были выделены ближайшие части границ обеих триангуляций, то необходимо подготовить их к объединению, чтобы в них было равное количество точек и они совпадали в пространстве. Для этого используется алгоритм корректировки ближайших границ. В ходе его работы из точек ближайших границ формируются ломаные линии. Изначально обе ломаные ориентируются друг относительно друга в пространстве так, чтобы их стартовые точки были близкими. Затем, начиная от стартовой точки одной ломаной происходит построение проекций ее узловых точек на вторую. Если проекция попадает на точку второй ломаной, или же находится вблизи этой точки, то такие точки будут являться близкими, значит их можно будет объединить в одну по среднему значению их координат. Если же проекция не попадает в близкую окрестность точки второй ломаной, то во вторую ломаную добавляется точка первой ломаной, от которой была построена проекция. Затем происходит построение проекций от точек второй ломаной на точки первой. В ходе работы этого алгоритма формируется список точек обеих триангуляций, которые будут являться близкими, и производится корректировка исходных точек обеих триангуляций с добавлением новых. Добавление новых точек в границу сопровождается разделением треугольника, содержащего пару граничных точек, между которыми добавляется новая точка, на два треугольника [4]. На рис. 4 показаны согласованные участки двух триангуляций из примера на рис. 3. В процессе согласования была добавлена одна вершина в первую триангуляцию, что повлекло за собой деление треугольника на два.

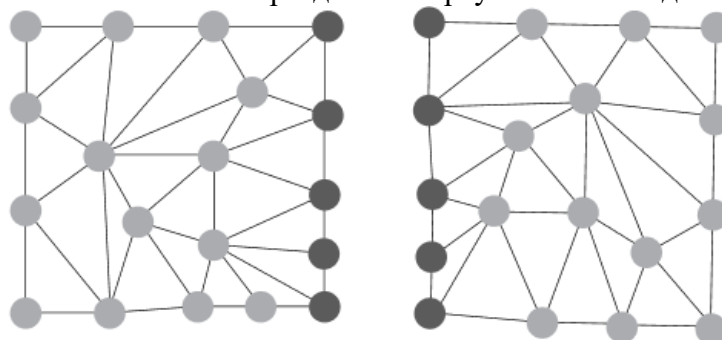


Рис. 4. Корректировка точек общих границ триангуляций

Сшивка триангуляций. После согласования границ обе триангуляции можно объединить в одну. Для этого последовательно осуществляется перенос в новую триангуляцию всех точек и треугольников первой триангуляции, а затем, используя индексы ближайших граничных точек первой и второй триангуляций, осуществляется перенос точек и треугольников второй триангуляции. Если переносится точка второй триангуляции, которая является близкой, то ей присваивается индекс близкой точки первой триангуляции, если же далекой, то присваивается новый индекс. Таким же образом производится корректировка индексов вершин треугольников второй триангуляции при их переносе.

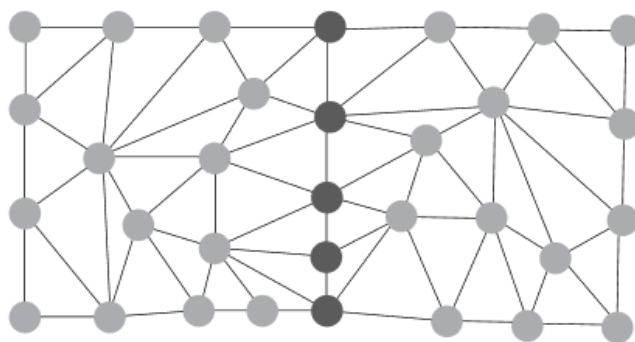


Рис. 5. Сшивание двух триангуляций в одну

Немаловажным фактором, который напрямую влияет на сшивку нескольких триангуляций, является отсутствие в них каких-либо изъянов. Из-за особенностей моделирования некоторых типов поверхностей и получения их сеточного представления одиночные триангуляции могут обладать близкими участками границы сами по себе (например, боковая поверхность цилиндра на границе области определения). Такие особенности напрямую влияют на работу алгоритма объединения, так как на месте шва происходит неправильное расположение треугольников, из-за чего существует возможность некорректного выделения границы такой триангуляции. Поэтому, для дальнейшей корректной работы алгоритма, необходимо осуществить проверку начальных триангуляций, которая выглядит следующим образом: изначально все точки данной триангуляции проверяются на уникальность, то есть не должно быть точек дубликатов, и на то, что они являются вершинами треугольников, висячих точек быть не должно (точек, которые не принадлежат ни одному треугольнику), в случае необходимости производится удаление таких точек и корректировка вершин треугольников. Затем происходит обработка шва построенной триангуляции в случае его наличия. Для этого определяются индексы начальных и конечных точек шва, по данным индексам точек разделяется граница триангуляции и получаются две ломаные, которые являются близкими.

Для тестирования и дальнейшей доработки алгоритма объединения нескольких триангуляций в единую был сформирован набор тестовых моделей с различным количеством точек для каждого из классов охватываемых им задач. Таким образом в тестах принимали участие как модели, состоящие из десятков точек, так и модели с десятками тысяч точек, результат работы алгоритма на которых был полностью верным, но время его работы отличалось. Следовательно, можно составить примерное соотношение времени работы данного алгоритма к задачам с различным количеством точек, но также стоит принимать во внимание и тот факт, что модели с одинаковым числом точек могут обрабатываться алгоритмом за различные промежутки времени. Связано это с тем, что начальные триангуляции могут иметь различные по числу точек общие границы, что повлечет за собой различия по времени в корректировке их исходных точек и треугольников, а далее и в их дальнейшем объединении с другими триангуляциями. График зависимости времени выполнения алгоритма от числа точек задачи изображен на рис. 6, оси графика представлены в логарифмическом масштабе.

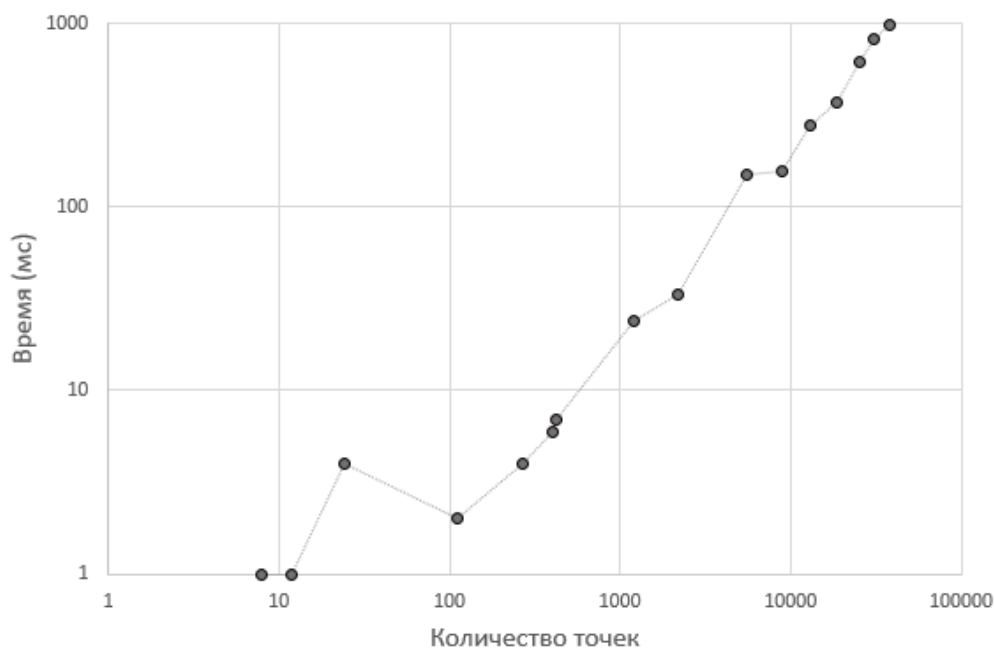


Рис. 6 .График зависимости размерности задачи от времени ее выполнения

Заключение

В статье описан алгоритм объединения нескольких триангуляций граней твердого тела в одну, были выделены основные его этапы и проблемы, которые возникали в ходе его работы на тестовых моделях.

Текущая реализация алгоритма показала приемлемую скорость работы при решении задач инженерного анализа. В дальнейшем планируется реализация параллельной версии данного алгоритма для оптимизации скорости его работы с моделями больших порядков.

Библиографический список

1. Данилов А.А. Способы построения трехмерных поверхностных триангуляций и тетраэдральных сеток // Системы автоматизированного проектирования // Научно-технический вестник Санкт-Петербургского государственного университета информационных технологий, механики и оптики. 2010. №1. С. 87-92.
2. Тюкачев Н.А. Моделирование поверхности трехмерных тел нерегулярной треугольной сеткой // Вестник Воронежского государственного университета. 2009. №1. С. 31-41.
3. Лебединская Н.А., Лебединский Д.М. Преобразование триангуляций при помощи элементарных операций // Вестник Санкт-Петербургского университета. 2009. №1. С. 84-86.
4. Тюкачев Н.А. Сгущение триангуляции поверхности со сглаживанием // Вестник Воронежского государственного университета. 2010. №2. С.33-38.

D.S. Kulikov, N.V. Starostin, A.E. Timofeev

**CONSTRUCTION OF A UNIFIED TRIANGULATION
OF A GEOMETRIC BODY SURFACE**

Lobachevsky State University of Nizhny Novgorod

Purpose: An important task in the subject field of engineering analysis is to work with the grid-based representation of geometric objects, in particular, solids. One of the topical problems is obtaining a grid-based representation of the surface of a geometric solid object with independent mesh representations (triangulations) of the solid faces.

Design/methodology/approach: The problem of combining grid representations tasks in intelligence work is considered in the case when grid representations are triangulations. It presents a serial algorithm for combining several triangulations into a single one. **Findings:** The main stages of this algorithm are identified and described. The operational analyses of the algorithm on practical models of different degrees of complexity is carried out.

Originality/value: The developed algorithm and program can be used to combine several mesh representations (triangulations) into a single one for engineering analysis tasks.

Keywords: grid-based representation, triangulation, engineering analysis.

ОПТИМАЛЬНОЕ УПРАВЛЕНИЕ ПРОДОЛЬНОМ ДВИЖЕНИЕМ БЕСПИЛОТНОГО ЛЕТАТЕЛЬНОГО АППАРАТА В УСЛОВИЯХ СМЕНЫ ВЫСОТ ПОЛЁТА

Арзамасский политехнический институт (филиал) НГТУ им. Р.Е.Алексеева

В статье рассматривается продольное движение беспилотного летательного аппарата в условиях смены высот полёта. Приводится решение задачи синтеза оптимального управления.

Ключевые слова: беспилотный летательный аппарат, система управления, траектория полёта, закон управления, оптимальное управление.

Конец XX – начало XXI века характеризуется тенденцией к расширению спектра применения беспилотных летательных аппаратов, от использования в гражданских нуждах [1] до задействования в условиях военных конфликтов [2]. Вне зависимости от области применения БПЛА, одной из наиболее важных задач является обеспечение устойчивости его движения в продольной плоскости. Технические требования на большинство летательных аппаратов (в том числе беспилотных) допускают их применение в определённом высотном диапазоне и подразумевают возможность изменения высоты полёта непосредственно в процессе траекторного движения. Для обеспечения безопасности и эффективности функционирования БПЛА в таких условиях необходимо организовать оптимальный [3] переходный процесс по высоте, исключающий возможность перерегулирования.

В общем случае необходимо обеспечить траекторию движения, аналогичную представленной на рис. 1. Здесь h_0 – начальная высота полёта летательного аппарата, $h_{зад}$ – требуемая высота полёта, $h_{зад} < h_0$ (смена высот при $h_{зад} > h_0$ выполняется аналогичным образом).

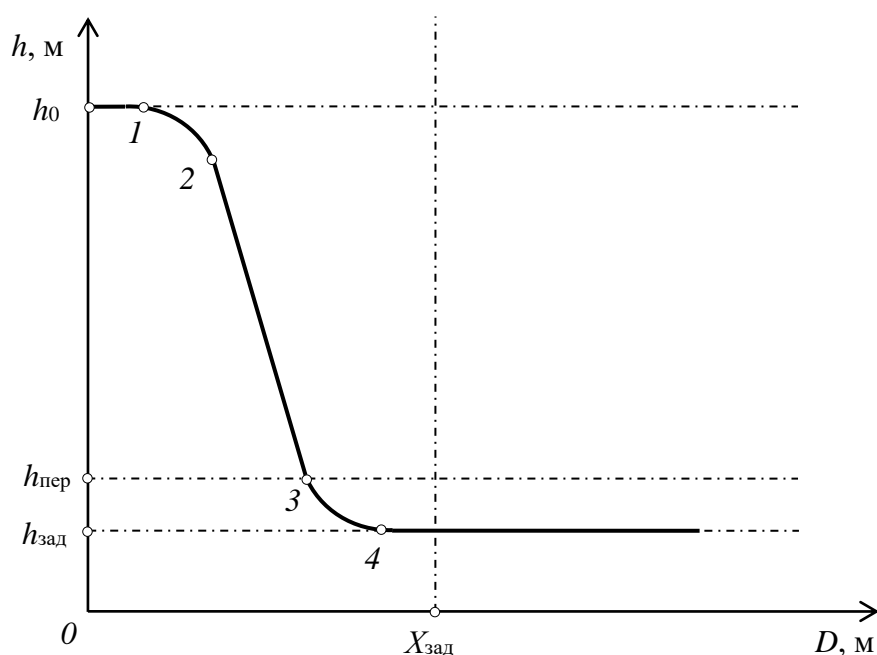


Рис. 1. Траектория движения БПЛА

Переходный процесс, представленный на рис. 1, включает в себя:

1. ввод БПЛА в пикирование (участок 1-2);
2. движение со стабилизацией постоянного угла тангажа $\theta_{зад}$ (участок 2-3);
3. вывод летательного аппарата из пикирования (участок 3-4) и дальнейшая стабилизация полёта на высоте $h_{зад}$.

Наибольшая эффективность алгоритма достигается за счёт обеспечения на участках 1-2 и 3-4 граничных значений управляющих перегрузок n_{y1} и n_{y2} :

$$n_{y1} \approx n_{y\max}, n_{y2} \approx -n_{y\max}, \quad (1)$$

где $n_{y\max}$ – максимальная по модулю перегрузка, сонаправленная с вектором ускорения свободного падения, ед. пер.

Итак, оптимальное управление имеет следующий вид:

$$u = n_y^* = \begin{cases} n_{y1}, & \text{(участок 1-2)} \\ \cos \theta_{зад}, & \text{(2-3)} \\ n_{y2} & \text{(3-4)} \end{cases} \quad (2)$$

Определим граничные условия для переключения управляющих сигналов. В полётном задании регламентируется координата $X_{зад}$ (рис. 1); переход на уровень высоты $h_{зад}$ должен быть обеспечен заблаговременно до достижения данной точки. Приведённое условие является основанием для выбора момента времени, соответствующего переходу БПЛА на участок пикирования.

Величина $\theta_{зад}$ определяется исходя из конструктивных особенностей измерителя высоты, используемого в составе системы управления летательного аппарата:

$$\theta \in [-\theta_{\max}, \theta_{\max}]. \quad (3)$$

Здесь θ_{\max} определяет диапазон изменения угла тангажа θ , при котором значения высоты, полученные при измерениях, будут иметь достаточную для практических расчётов точность.

Переход в режим стабилизации угла тангажа, соответствующий участку траекторного движения 2-3, выполняется при достижении граничного значения θ_{\max} :

$$|\theta| = \theta_{\max}, \quad (4)$$

соответственно,

$$\theta_{зад} = \theta_{\max}. \quad (5)$$

Момент начала вывода беспилотного летательного аппарата из пикирования определяется достижением в процессе полёта рассчитанной высоты $h_{пер}$ (точка 3 на рисунке 1). Её вычисление производится в соответствии с условием безопасности полёта:

$$\Delta h < 0,05 \cdot h_{зад}. \quad (6)$$

где Δh – абсолютная величина перерегулирования [4] по высоте, м.

Выполним расчёт величины $h_{пер}$. Продольному движению БПЛА соответствует система уравнений:

$$\begin{cases} \dot{\theta} = \frac{(n_y - \cos \theta)g}{V}, \\ \dot{h} = V \cdot \sin \theta, \\ \dot{V} = g [n_x (h, V, n_y, \xi) - \sin \theta]. \end{cases} \quad (7)$$

где n_y – нормальная перегрузка, ед. пер; g – ускорение свободного падения, м/с²; V – скорость движения ЛА, м/с; h – высота полёта, м; n_x – продольная перегрузка, ед. пер; ξ – неопределенный параметр с известным диапазоном изменения.

Допущение о постоянстве скорости на всём протяжении переходного процесса позволяет упростить исходную систему уравнений:

$$\begin{cases} \dot{\theta} = \frac{(n_y - \cos \theta)g}{V_c}, \\ \dot{h} = V_c \cdot \sin \theta \end{cases}, \quad V_c = \text{const}. \quad (8)$$

Для получения решения данной системы воспользуемся принципом максимума Понтрягина [5]. Гамильтониан системы (8) имеет вид:

$$H = \psi_1 \cdot \frac{(n_y - \cos \theta)g}{V_c} + \psi_2 \cdot V_c \cdot \sin \theta, \quad (9)$$

где

$$\begin{cases} \dot{\psi}_1 = -\frac{\partial H}{\partial \theta} = -\psi_1 \cdot \frac{g \sin \theta}{V_c} - \psi_2 \cdot V_c \cdot \cos \theta, \\ \dot{\psi}_2 = -\frac{\partial H}{\partial h} = 0. \end{cases} \quad (10)$$

Экстремум приведённой функции Гамильтона (9) достигается при условии

$$n_y^* = n_{y \max} \cdot \text{sign}(\psi_1). \quad (11)$$

Интегрирование уравнений системы (7) в соответствии с равенством (10) позволяет определить величину высоты $h_{\text{пер}}$, при которой необходимо выполнить переключение управляющего сигнала:

$$h_{\text{пер}} = h_0 + \ln \left(\frac{n_{y \max} - \cos \theta}{n_{y \max} - \cos \theta_0} \right) \cdot \frac{V_c^2}{g}. \quad (12)$$

Учитывая, что движение БПЛА на начальном участке происходит в режиме стабилизации заданной высоты H_0 , начальное значение угла тангажа $\theta_0 \rightarrow 0$.

С учётом (11) управление (2) примет вид:

$$n_y = \begin{cases} \left. \begin{matrix} n_{y1}, & |\theta| < \theta_{max}, \\ \cos \theta, & |\theta| = \theta_{max} \end{matrix} \right\} & h > h_0 + \ln \left(\frac{n_{y_{max}} - \cos \theta}{n_{y_{max}} - 1} \right) \cdot \frac{V_c^2}{g}, \\ n_{y2}, & h = h_0 + \ln \left(\frac{n_{y_{max}} - \cos \theta}{n_{y_{max}} - 1} \right) \cdot \frac{V_c^2}{g}. \end{cases} \quad (13)$$

Полученный результат в дальнейшем может быть применен для формирования алгоритма продольного движения БПЛА в условиях смены высот.

Библиографический список

1. Беспилотный летательный аппарат [Электронный ресурс] / Государство. Бизнес. ИТ. URL: [http://www.tadviser.ru/index.php/Статья:Беспилотный_летательный_аппарат_\(дрон,_БПЛА\)](http://www.tadviser.ru/index.php/Статья:Беспилотный_летательный_аппарат_(дрон,_БПЛА))
2. Российские и зарубежные беспилотники [Электронный ресурс] / Военный обзор. URL: <https://militaryarms.ru/voennaya-texnika/aviaciya/bespilotnye-letatelnye-apparaty>
3. Математическая теория оптимальных процессов / Л. С. Понтрягин, В.Г. Болтянский, Р.В. Гамкрелидзе, Е.Ф. Мищенко. – М.: Государственное издательство физико-математической литературы, 1961. – 394 с.
4. Теория систем автоматического регулирования / В. А. Бесекерский, Е. П. Попов. – СПб.: Профессия, 2004. — 749 с.
5. Понтрягин Л. С. Принцип максимума в оптимальном управлении / Л. С. Понтрягин. – М.: Главная редакция физико-математической литературы издательства "Наука", 1989. – 64 с.

N. D. Khaleev

OPTIMAL LONGITUDINAL MOTION CONTROL OF UNMANNED AERIAL VEHICLE UNDER CHANGING OF THE FLIGHT ALTITUDE

Nizhny Novgorod state technical university n a. R.E. Alekseev

The paper considers longitudinal motion of an unmanned aerial vehicle changing of the flight altitudes. The solution to the problem of optimal control design is presented.

Keywords: unmanned aerial vehicle, control system, flight trajectory, control law, optimal control.

УПРАВЛЕНИЕ С ИТЕРАТИВНЫМ ОБУЧЕНИЕМ ГРУППОЙ СИСТЕМ С НЕОПРЕДЕЛЕННЫМИ ПАРАМЕТРАМИ ПРИ ИЗМЕНЕНИИ КОНФИГУРАЦИИ ИНФОРМАЦИОННОЙ СЕТИ

Нижегородский государственный технический университет им. Р.Е. Алексеева

В статье рассматривается задача синтеза управления с итеративным обучением группой систем (агентов) с неопределенными параметрами, связанных информационной сетью. Агенты представляют собой линейные дискретные системы с аффинными моделями неопределенностей, работающие в повторяющемся режиме и могут подключаться к сети и отключаться от нее в соответствии с определенными внешними правилами. Синтез управления основан на дивергентном методе векторной функции Ляпунова. Приводятся результаты моделирования полученного закона управления с итеративным обучением группой однозвенных гибких манипуляторов при последовательном подключении агентов.

Ключевые слова: сетевое управление, управление с итеративным обучением, системы с неопределенными параметрами, системы с переключениями, повторяющиеся процессы, 2D-системы, устойчивость, векторная функция Ляпунова.

Введение

Системы с переключениями – это класс динамических систем, состоящих из конечного числа подсистем и некоторого логического правила, обеспечивающего переключение между ними. Данный класс систем интенсивно изучается в последние десятилетия исследователями различных профилей, в первую очередь специалистами по теории управления, компьютерным наукам и прикладной математике. Это объясняется тем фактом, что многие инженерные системы могут моделироваться с использованием такой структуры.

Целью данной работы является исследование управления сетевыми системами с переключениями. В общем случае под сетевыми системами понимается совокупность подсистем (агентов), между которыми установлены физические или информационные связи. Примерами таких систем являются транспортные и высокотехнологичные производственные сети, а также системы координированного управления групповым движением роботов.

К системам, функционирующим в повторяющемся режиме, эффективно применяется метод управления с итеративным обучением (УИО). С момента первой публикации [1], управление с итеративным обучением является интенсивно развивающейся областью исследований. Метод УИО базируется на идее о том, что точность работы системы, многократно выполняющей одну и ту же операцию, может быть улучшена за счет результатов с предыдущих повторений. Управление с итеративным обучением позволяет достичь требуемой точности при наличии динамики, точное моделирование которой не представляется возможным, или параметрических неопределенностей, проявляющихся во время фактической работы системы.

В данной работе рассматривается управление с итеративным обучением группой систем с неопределенными параметрами, связанных информационной сетью. Агенты могут подключаться к сети и отключаться от нее в соответствии с определенными внешними правилами. При этом всякий раз меняется конфигурация сети, т. е. количество агентов в сети и характер обмена информацией. Агенты сетевой системы представляют собой линейные системы с дискретным временем и аффинными моделями

параметрических неопределенностей. Синтез закона УИО осуществляется на основе дивергентного метода векторных функций Ляпунова [2], который получил развитие для систем с переключениями в работе [3].

Описание сетевой системы и постановка задачи

Рассмотрим систему из N линейных динамических систем (агентов) с неопределенными параметрами, функционирующих в повторяющемся режиме и имеющих возможность подключаться и к сети отключаться от нее. Конфигурация m сетевой системы определяет группу из N_m агентов, функционирующих на данной конфигурации. Динамика агента i , подключенного к сети в конфигурации m на повторении k , описывается следующей дискретной моделью в пространстве состояний:

$$\begin{aligned} x_i(p+1, k) &= A(\delta_i(p))x_i(p, k) + B(\delta_i(p))u_i(p, k), \\ y_i(p, k) &= Cx_i(p, k), \quad i \in J_m, \quad k \geq 0, \quad 0 \leq p \leq T-1, \end{aligned} \quad 1)$$

с граничными условиями $x_i(0, k) = x_0$, $y_i(p, 0) = Cx_i(p, 0)$ и $u_i(p, 0) = 0$, где $x_i(p, k) \in \mathbb{R}^{n_x}$ – вектор состояния агента i , $u_i(p, k) \in \mathbb{R}^{n_u}$ – вектор управления агента i , $y_i(p, k) \in \mathbb{R}^{n_y}$ – вектор выходных переменных (профиль повторения) агента i , x_0 – известные постоянные, $J_m = \{i_1, i_2, \dots, i_{N_m}\}$ – набор номеров агентов, доступных в конфигурации m , $T < \infty$ – продолжительность повторения.

Модель неопределенностей задается в виде

$$\begin{aligned} A(\delta_i(p)) &= A + \sum_{n=1}^l \delta_{i,n}(p)A_n, \\ B(\delta_i(p)) &= B + \sum_{n=1}^l \delta_{i,n}(p)B_n, \end{aligned}$$

где A и B – матрицы номинальной модели, A_n и B_n ($n = 1, 2, \dots, l$) – постоянные матрицы соответствующих размеров и $\delta_{i,n}(p) \in [\underline{\delta}_{i,n}, \bar{\delta}_{i,n}]$.

Обозначим

$$\begin{aligned} \Delta_i &= \{\delta_i = [\delta_{i,1} \ \dots \ \delta_{i,l}]^T : \delta_{i,n} \in [\underline{\delta}_{i,n}, \bar{\delta}_{i,n}], n = 1, 2, \dots, l\}, \\ \Delta_{v,i} &= \{\delta_i = [\delta_{i,1} \ \dots \ \delta_{i,l}]^T : \delta_{i,n} \in \{\underline{\delta}_{i,n}, \bar{\delta}_{i,n}\}, n = 1, 2, \dots, l\}, \end{aligned}$$

где Δ_i – множество неопределенных параметров агента i , $\Delta_{v,i}$ – конечное множество вершин многогранника ограничений агента i , содержащее 2^l элементов.

Связи между агентами в конфигурации m представим в виде направленного графа $\mathcal{G}_m = (J_m, \mathcal{E}_m)$, где $\mathcal{E}_m \subseteq J_m \times J_m$ – ребра графа. Возможность доступа агента $j \in J_m$ к выходным данным агента i задается ребром, направленным от вершины i к вершине j и обозначаемым упорядоченной парой $(i, j) \in \mathcal{E}_m$. Элементы матрицы смежности

$$D_m(\mathcal{G}_m) = \begin{bmatrix} d_{i_1 i_1} & d_{i_1 i_2} & \dots & d_{i_1 i_{N_m}} \\ d_{i_2 i_1} & d_{i_2 i_2} & \dots & d_{i_2 i_{N_m}} \\ \vdots & \vdots & \ddots & \vdots \\ d_{i_{N_m} i_1} & d_{i_{N_m} i_2} & \dots & d_{i_{N_m} i_{N_m}} \end{bmatrix}$$

задаются следующим образом: $d_{ij} > 0$ если $(j, i) \in \mathcal{E}_m$, $d_{ij} = 0$ в противном случае,

$d_{ii} = 0$. Значения ненулевых элементов, в общем случае, различные и отражают значимость тех или иных связей. В простейшем случае они равны единице.

Матрица Лапласа графа \mathcal{G}_m задается выражением

$$L_m(\mathcal{G}_m) = \begin{bmatrix} \sum_{j \in J_m} d_{i_1 j} & -d_{i_1 i_2} & \dots & -d_{i_1 i_{N_m}} \\ -d_{i_2 i_1} & \sum_{j \in J_m} d_{i_2 j} & \dots & -d_{i_2 i_{N_m}} \\ \vdots & \vdots & \ddots & \vdots \\ -d_{i_{N_m} i_1} & -d_{i_{N_m} i_2} & \dots & \sum_{j \in J_m} d_{i_{N_m} j} \end{bmatrix}.$$

Обозначим за $y_{ref}(p)$ желаемый выходной сигнал и введем ошибку выходного сигнала $e_i(p, k) = y_{ref}(p) - y_i(p, k)$. В рамках задачи сетевого управления, агента,

имеющего доступ к $y_{ref}(p)$, будем называть лидером, а остальных – ведомыми системами. Возможность получения агентами информации о желаемой траектории задается матрицей

$$G_m = \begin{bmatrix} g_{i_1} & 0 & \dots & 0 \\ 0 & g_{i_2} & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & g_{i_{N_m}} \end{bmatrix},$$

где $g_i = 1$ если агенту i имеет доступ к $y_{ref}(p)$, и $g_i = 0$ в противном случае.

Задача заключается в поиске такого управления $u_i(p, k)$, при котором производительность постепенно улучшается с каждым повторением. Это можно обобщить следующими условиями сходимости ошибки и входного сигнала

$$\lim_{k \rightarrow \infty} |e_i(p, k)| = 0, \quad \lim_{k \rightarrow \infty} |u_i(p, k) - u_i(p, \infty)| = 0,$$

где переменная $u_i(p, \infty)$ ограничена и называется обученным управлением.

Закон УИО для агента i на повторении $k + 1$ в случае, если он функционировал на повторении k имеет вид

$$u_i(p, k + 1) = u_i(p, k) + \Delta u_i(p, k + 1), \quad (2)$$

где $\Delta u_i(p, k + 1)$ – корректирующая добавка для агента i , которую будем формировать в виде

$$\Delta u_i(p, k + 1) = K_{1,m} [x_i(p, k + 1) - x_i(p, k)] + K_{2,m} \left[\sum_{j \in \mathcal{J}_{i,m}} d_{ij} (e_i(p + 1, k) - e_j(p + 1, k)) + g_i e_i(p + 1, k) \right]. \quad (3)$$

где $K_{1,m}$ и $K_{2,m}$ – матрицы протокола в конфигурации m , $\mathcal{J}_{i,m} = \{j \in \mathcal{J}_m | (j, i) \in \mathcal{E}_m\}$ – множество агентов, информация о которых доступна агенту i в этой конфигурации.

Если агент i подключается к сети на повторении $k + 1$, то он получает управляющий сигнал от агента j при условии, что агент j функционировал на повторении k и продолжает свою работу на повторении $k + 1$, т. е.

$$u_i(p, k + 1) = u_j(p, k) + \Delta u_j(p, k + 1). \quad (4)$$

Конфигурацию сетевой системы, при которой управление хотя бы одним агентом происходит согласно закону (4), будем называть переходной конфигурацией, в противном случае – стандартной конфигурацией.

Построение 2D-модели

Введем вектор $\xi_i(p + 1, k + 1) = x_i(p, k + 1) - x_i(p, k)$, расширенные векторы

$$\xi(p, k) = [\xi_{i_1}^T(p, k) \quad \dots \quad \xi_{i_{N_m}}^T(p, k)]^T, \quad e(p, k) = [e_{i_1}^T(p, k) \quad \dots \quad e_{i_{N_m}}^T(p, k)]^T,$$

и запишем динамику управления с итеративным обучением как дискретный линейный повторяющийся процесс относительно расширенных векторов:

$$\xi(p + 1, k + 1) = \bar{A}_m(\bar{\delta}_m) \xi(p, k + 1) + \bar{B}_m(\bar{\delta}_m) \bar{H}_m v(p, k + 1),$$

$$e(p, k + 1) = -\bar{C}_m \bar{A}_m(\bar{\delta}_m) \xi(p, k + 1) + e(p, k) - \bar{C}_m \bar{B}_m(\bar{\delta}_m) \bar{H}_m v(p, k + 1),$$

где $v(p, k + 1) = \bar{K}_{1,m} \xi(p, k + 1) + \bar{K}_{2,m} \bar{\Gamma}_m e(p, k)$ – корректирующая добавка,

$$\bar{A}_m(\bar{\delta}_m) = \text{diag} [A(\delta_{i_1}) \quad \dots \quad A(\delta_{i_{N_m}})], \quad \bar{B}_m(\bar{\delta}_m) = \text{diag} [B(\delta_{i_1}) \quad \dots \quad B(\delta_{i_{N_m}})],$$

$$\bar{C}_m = I_{N_m} \otimes C, \quad \bar{K}_{1,m} = I_{N_m} \otimes K_{1,m}, \quad \bar{K}_{2,m} = I_{N_m} \otimes K_{2,m},$$

$$\bar{H}_m = H_m \otimes I_{n_u}, \quad \bar{\Gamma}_m = (L_m + G_m) \otimes I_{n_y}.$$

Здесь $\bar{\delta}_m = [\delta_{i_1}^T \quad \dots \quad \delta_{i_{N_m}}^T]^T$ – расширенный вектор неопределенных параметров.

При этом множество неопределенных параметров и конечное множество его 2^{LN_m} вершин определяются как

$$\bar{\Delta}_m = \left\{ \bar{\delta}_m = [\delta_{i_1}^T \quad \dots \quad \delta_{i_{N_m}}^T]^T : \delta_i \in \Delta_i, i \in \mathcal{J}_m \right\},$$

$$\bar{\Delta}_{v,m} = \left\{ \bar{\delta}_m = [\delta_{i_1}^T \quad \dots \quad \delta_{i_{N_m}}^T]^T : \delta_i \in \Delta_{v,i}, i \in \mathcal{J}_m \right\}.$$

Элементы матрицы

$$H_m = \begin{bmatrix} h_{i_1 i_1} & h_{i_1 i_2} & \dots & h_{i_1 i_{N_m}} \\ h_{i_2 i_1} & h_{i_2 i_2} & \dots & h_{i_2 i_{N_m}} \\ \vdots & \vdots & \ddots & \vdots \\ h_{i_{N_m} i_1} & h_{i_{N_m} i_2} & \dots & h_{i_{N_m} i_{N_m}} \end{bmatrix},$$

задаются следующим образом: $h_{ij} = 1$ если на режиме m агент j передает агенту i свой управляющий сигнал, в противном случае $h_{ij} = 0$; $h_{ii} = 1$ если на режиме m агент i работает самостоятельно, в противном случае $h_{ii} = 0$.

Решение на основе дивергентного метода векторных функций Ляпунова

При заданной структуре управления задача синтеза управления сводится к нахождению матриц протокола $K_{1,m}$ и $K_{2,m}$. В результате применения дивергентного метода векторной функции Ляпунова [3] получим, что эти матрицы удовлетворяют следующей системе матричных уравнений и неравенств:

$$\begin{bmatrix} X_m & (A_m(\bar{\delta}_m)X_m + B_m(\bar{\delta}_m)Y_m\Gamma_m)^T & X_m & (Y_m\Gamma_m)^T \\ A_m(\bar{\delta}_m)X_m + B_m(\bar{\delta}_m)Y_m\Gamma_m & X_m & 0 & 0 \\ X_m & 0 & Q_m^{-1} & 0 \\ Y_m\Gamma_m & 0 & 0 & R_m^{-1} \end{bmatrix} \succcurlyeq 0 \quad (5)$$

$$\Gamma_m X_m = Z_m \Gamma_m, X_m > 0,$$

где $Q_m > 0, R_m > 0$ – весовые матрицы,

$$A_m(\bar{\delta}_m) = \begin{bmatrix} \bar{A}_m(\bar{\delta}_m) & 0 \\ -\bar{C}_m \bar{A}_m(\bar{\delta}_m) & I_{N_m} \otimes I_{n_y} \end{bmatrix}, B_m(\bar{\delta}_m) = \begin{bmatrix} \bar{B}_m(\bar{\delta}_m) \bar{H}_m \\ -\bar{C}_m \bar{B}_m(\bar{\delta}_m) \bar{H}_m \end{bmatrix},$$

$$\Gamma = \begin{bmatrix} I_{N_m} \otimes I_{n_x} & 0 \\ 0 & \bar{\Gamma}_m \end{bmatrix},$$

$$X_m = \text{diag}[\bar{X}_{1,m}, \bar{X}_{2,m}], \bar{X}_{1,m} = I_{N_m} \otimes X_{1,m}, \bar{X}_{2,m} = I_{N_m} \otimes X_{2,m},$$

$$Y_m = [\bar{Y}_{1,m}, \bar{Y}_{2,m}] = [\bar{K}_{1,m}, \bar{K}_{2,m}] Z_m, \bar{Y}_{1,m} = I_{N_m} \otimes Y_{1,m}, \bar{Y}_{2,m} = I_{N_m} \otimes Y_{2,m},$$

$$Z_m = \text{diag}[\bar{Z}_{1,m}, \bar{Z}_{2,m}], \bar{Z}_{1,m} = I_{N_m} \otimes Z_{1,m}, \bar{Z}_{2,m} = I_{N_m} \otimes Z_{2,m},$$

$$Q_m = \text{diag}[\bar{Q}_{1,m}, \bar{Q}_{2,m}], \bar{Q}_{1,m} = I_{N_m} \otimes Q_1, \bar{Q}_{2,m} = I_{N_m} \otimes Q_2, R_m = I_{N_m} \otimes R.$$

Если система (5) разрешима относительно $X_{1,m}, X_{2,m}, Y_{1,m}, Y_{2,m}, Z_{1,m}$ и $Z_{2,m}$ для $\bar{\delta}_m \in \bar{\Delta}_{v,m}$, то закон управления (2), (4) с корректирующей добавкой (3) и матрицами протокола $K_{1,m} = Y_{1,m}(Z_{1,m})^{-1}$ и $K_{2,m} = K_{2,m}(Z_{2,m})^{-1}$ гарантирует экспоненциальную устойчивость системы и обеспечит сходимость выходного сигнала к желаемой траектории.

Пример

В качестве агента сетевой системы будем рассматривать манипулятор с одним гибким звеном [4], момент инерции которого точно не известен и задан в виде аффинной модели. Динамика движения манипулятора в этом случае описывается следующей моделью в пространстве состояний:

$$\begin{aligned} \dot{x} &= A_c(\delta)x + B_c u, \\ y &= C_c x, \end{aligned}$$

где $x = [\theta \ \alpha \ \dot{\theta} \ \dot{\alpha}]^T$, θ – угол поворота сервопривода, α – угол отклонения звена,

$u = \tau$ – крутящий момент, приложенный к основанию гибкого звена,

$$A_c(\delta) = \begin{bmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & \frac{K_s}{J_{\varepsilon q}} & -\frac{B_{\varepsilon q}}{J_{\varepsilon q}} & 0 \\ 0 & -\frac{K_s}{J_{\varepsilon q}} & \frac{B_{\varepsilon q}}{J_{\varepsilon q}} & 0 \end{bmatrix} + \delta_1 \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & -\frac{K_s}{J_l} & 0 & 0 \end{bmatrix}, B_c = \begin{bmatrix} 0 \\ 0 \\ \frac{1}{J_{\varepsilon q}} \\ -\frac{1}{J_{\varepsilon q}} \end{bmatrix},$$

$$C_c = [1 \ 0 \ 0 \ 0],$$

K_s – жесткость гибкого звена, $J_{\varepsilon q}$ – момент инерции сервопривода, $B_{\varepsilon q}$ – коэффициент вязкого трения сервопривода, J_l – момент инерции звена относительно центра масс, $\delta_1 \in [\underline{\delta}_1, \bar{\delta}_1]$.

Рассмотрим случай, когда система состоит из трех одинаковых агентов, лидер получает информацию о желаемой траектории, а ведомые системы получают информацию о выходном сигнале от лидера. Сценарий работы системы определяется следующими конфигурациями:

1. работает только лидер;
2. подключается первая ведомая система и получает управляющий сигнал от лидера;
3. лидер и первая ведомая система работают самостоятельно, получая доступную им информацию о выходном сигнале;
4. подключается вторая ведомая система и получает управляющий сигнал от лидера;
5. все агенты работают самостоятельно, получая доступную им информацию о выходном сигнале.

Такая модель соответствует производственной системе с изменением объема производства. Когда объем производства увеличивается, происходит подключение новых агентов, при уменьшении – отключение лишних. С целью снижения нагрузки на сеть, сетевая система работает на переходных режимах 2 и 4 не более одного повторения.

Были приняты следующие значения параметров однозвенных гибких манипуляторов:

$K_s = 1,3 \text{ Н} \cdot \text{м/рад}$, $J_{\varepsilon q} = 2,08 \times 10^{-3} \text{ кг} \cdot \text{м}^2$, $B_{\varepsilon q} = 0,004 \text{ Н} \cdot \text{м/(рад/с)}$,
 $J_l = 0,0038 \text{ кг} \cdot \text{м}^2$. Продолжительность цикла повторения составляет 3 с, требуемая точность $e^* = 0,01$ рад. При переходе к дискретной модели использовался метод Эйлера с шагом $T_s = 0,01$ с. Желаемая траектория выходного сигнала описывается уравнением

$$y_{ref}(p) = \frac{\pi(pT_s)^2}{6} - \frac{\pi(pT_s)^3}{27}, \quad 0 \leq p \leq T - 1, \quad T = 300.$$

В данном примере $CB = 0$, однако разработанная теория предполагает, что $CB \neq 0$. Для обхода этого затруднения воспользуемся тем, что для данной системы $\dot{y} = C_c \dot{x} = C_{c,1} x$, где $C_{c,1} = [0 \ 0 \ 1 \ 0]$. Тогда используя метод Эйлера, для решения задачи синтеза будем использовать следующую дискретную модель в пространстве состояний, описывающую динамику агента i на повторении k :

$$\begin{aligned} x_i(p+1, k) &= A(\delta_i)x_i(p, k) + Bu_i(p, k), \\ y_i(p+1, k) &= Cx_i(p, k), \end{aligned}$$

с граничными условиями $x_i(0, k) = x_0$, $y_i(0, k) = 0$, $u_i(p, 0) = 0$, где $A(\delta_i) = I_{n_x} + A_c(\delta_i)T_s$, $B = B_c T_s$, $C = C_c + C_{c,1}T_s$.

В данном случае введем ошибку обучения как $\epsilon_i(p, k) = y_{ref}(p+1) - y_i(p+1, k)$ и будем формировать закон управления в виде (2), (4) с корректирующей добавкой в виде

$$\begin{aligned} \Delta u_i(p, k+1) &= K_{1,m}[x_i(p, k+1) - x_i(p, k)] + \\ &+ K_{2,m}[\sum_{j \in J_{i,m}} d_{ij}(\epsilon_i(p+1, k) - \epsilon_i(p, k)) + g_i \epsilon_i(p+1, k)]. \end{aligned}$$

Тогда динамика управления с итеративным обучением будет описываться уравнениями

$$\xi(p+1, k+1) = \bar{A}_m(\bar{\delta}_m)\xi(p, k+1) + \bar{B}\bar{H}_m v(p, k+1),$$

$$\epsilon(p, k + 1) = -\bar{C}_m \bar{A}_m(\bar{\delta}_m) \xi(p, k + 1) + \epsilon(p, k) - \bar{C} \bar{B} \bar{H}_m v(p, k + 1),$$

где $v(p, k + 1) = \bar{K}_{1,m} \xi(p, k + 1) + \bar{K}_{2,m} \bar{\Gamma}_m \epsilon(p, k)$ – корректирующая добавка.

Решая задачу максимизации $\text{tr}[X_m]$ при ограничениях в виде ЛМН (5), задавая весовые матрицы $Q_1 = 10^{-6} I_{n_x}$, $Q_2 = 10^4$, $R = 10^3$, и изменяя матрицу $A_m(\bar{\delta}_m)$ относительно

неопределенных параметров агентов $\delta_{1,1} \in \{0,8; 1,2\}$, $\delta_{2,1} = \{0,79; 1,18\}$, $\delta_{3,1} \in \{0,82; 1,21\}$, были получены следующие матрицы протокола:

$$\begin{aligned} K_{1,1} &= [-14,951 \quad -0,7728 \quad -0,3717 \quad -0,0339], & K_{2,1} &= [9,7491], \\ K_{1,2} &= [-0,0573 \quad -0,0325 \quad -0,0073 \quad -0,0029], & K_{2,2} &= [0,0009], \\ K_{1,3} &= [-19,727 \quad -1,1769 \quad -0,4048 \quad -0,0061], & K_{2,3} &= [0,2733], \\ K_{1,4} &= [-0,0931 \quad -0,0559 \quad -0,0116 \quad -0,0045], & K_{2,4} &= [0,0012], \\ K_{1,5} &= [-19,719 \quad -1,163 \quad -0,4058 \quad -0,007], & K_{2,5} &= [0,2777]. \end{aligned}$$

На рисунке 1 представлен график изменения среднеквадратических ошибок выходных сигналов агентов в логарифмическом масштабе. Неопределенные параметры агентов находятся на верхних границах. Лидер достигает требуемую точность за 5 повторений, после чего происходит подключение ведомых систем на повторении 11 и 23 соответственно. Ошибки ведомых систем не превышают требуемого значения в **0,01** рад на протяжении всего времени функционирования. Вторая ведомая система отключается на повторении 81.

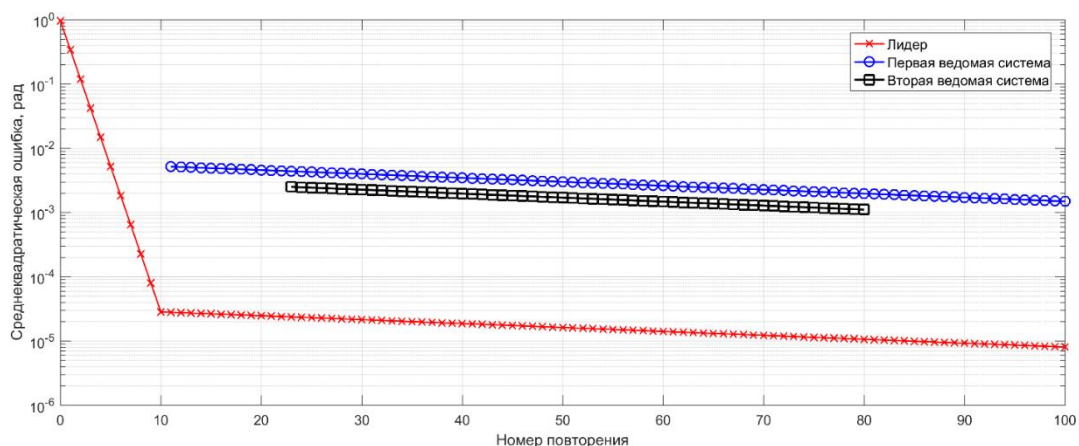


Рис. 1. Среднеквадратические ошибки выходных сигналов агентов

Заключение

В данной работе с использованием дивергентного метода векторной функции Ляпунова получен закон управления с итеративным обучением группой систем с неопределенными параметрами, связанных информационной сетью переменной конфигурации. Применение этого закона управления позволяет достичь высокой точности отслеживания заданной траектории движения в повторяющемся режиме, что подтверждается моделированием полученного алгоритма к группе из однозвенных гибких манипуляторов. В дальнейшем планируется развить дивергентный метод векторной функции Ляпунова для других классов сетевых систем с переключениями, в частности при учете случайных возмущений, действующих на объект, и шумов измерений.

Библиографический список

1. Arimoto S., Kawamura S., Miyazaki F. Bettering Operation of Robots by Learning // J. Robot. Syst. 1984. V. 1. P. 123–140.
2. Галковский К., Емельянов М. А., Пакшин П.В., Роджерс Э. Векторные функции Ляпунова в задачах устойчивости и стабилизации дифференциальных повторяющихся процессов // Изв. РАН. Теория и системы управления. 2016. №4. С. 5 – 17.
3. Pakshin P., Emelianova J. Stability of Nonlinear Discrete Repetitive Processes with Switching // 2020 European Control Conference (ECC), Saint Petersburg, Russia, 2020. P. 17–22.
4. Apkarian J., Karam P., L'evis M. Workbook on Flexible Link Experiment for Matlab/Simulink Users. Quanser, 2011.

A.S. Kopusov, P.V. Pakshin

ITERATIVE LEARNING CONTROL OF A GROUP OF SYSTEMS WITH UNCERTAIN PARAMETERS UNDER CHANGING THE INFORMATION NETWORK CONFIGURATION

Nizhny Novgorod state technical university n a. R.E. Alekseev

The paper considers the iterative learning control design problem of a group of systems (agents) with uncertain parameters connected by an information network. Agents are described by linear discrete-time models and uncertainties are given by affine models of uncertainty. They are operating in a repetitive model and can connect to the network and disconnect from it in accordance with certain external rules. The control design is based on the divergent method of the vector Lyapunov function. The results of modeling the obtained control law with iterative learning by a group of single-link flexible manipulators with sequential connection of agents are presented.

Keywords: networked control, control with iterative learning, systems with uncertain parameters, systems with switches, repetitive processes, 2D systems, stability, vector Lyapunov function.

ИЕРАРХИЯ КОНЦЕПЦИЙ И МОДЕЛЕЙ ПОСТРОЕНИЯ АВТОМАТИЗИРОВАННЫХ СИСТЕМ

Нижегородский государственный технический университет им. Р.Е. Алексеева

В статье описана иерархическая система концепций построения автоматизированных систем. Все концепции разделены на три уровня иерархии. На верхнем уровне классификационной системы расположены концепции, которые описывают глобальные тренды развития мира информационных технологий. На втором уровне – концепции, которые ориентированы на процесс проектирования автоматизированных систем. На самом низком уровне – третьем расположены концепции построения конкретных видов обеспечений автоматизированных систем.

Так же в статье рассматривается классификационная система моделей поддержки процесса проектирования АС. Эта система имеет два уровня иерархии. На высшем уровне расположены декларативные модели описания самого процесса, которые применимы для любого конкретного проекта. На низшем уровне расположены модели, которые проектно-ориентированы и поддерживают конкретные проектные операции и проектные процедуры.

Ключевые слова: концепция, иерархия, модели, фреймовая модель, фреймово-продукционная модель, фреймы с мультимедийными образами, системы управления, системы мониторинга, автоматизированные системы, проектирование.

Введение

Активное развитие автоматизированных систем, захват ими всех сфер жизни человеческого инициирует развитие теоретических основ их создания и поддержки в течение всего жизненного цикла. Наличие множества концепций и моделей актуализирует их анализ с целью формализации факторов успеха в комплексе вопросов их применения.

Отметим, что любая классификация концепций и моделей несет некоторый оттенок субъективизма. Однако, создание классификационных систем является научным методом исследования, который прямо или опосредованно приводит к часто весьма интересным результатам. В данной работе авторами предлагается классификационная система, которая охватывает концепции и модели построения автоматизированных систем.

Иерархия концепций

В данной работе предлагается условно разделить множество концепций на три уровня. К первому уровню (рис. 1) отнесем концепции, которые выступают в качестве общих устойчивых трендов развития ИТ мира. К ним отнесем:

- интеллектуализацию автоматизированных систем;
- глобализацию;
- использование системного подхода [1];
- поддержку системы в течение всего жизненного цикла [2];
- открытые системы [3];
- объектно-ориентированный подход [4];
- и другие.



Рис. 1. Концепции первого уровня иерархии

Ко второму уровню (рис.2) отнесем концепции, которые поддерживают сам процесс проектирования. Это концепции:

- ситуационного подхода [5];
- сценарного подхода [5];
- сценарно-ситуационного подхода [5];
- логистический подход (каждый факт, каждое данное в нужном формате в нужном месте и в нужный момент времени) [6];
- агентного подхода [7];
- и другие.

Условно к третьему уровню (рис.2) иерархии отнесем концепции, которые выступают в роли руководства к практическим действиям. Это в первую очередь концепции построения программного и информационного обеспечения информационных систем.

При построении информационного обеспечения используются концепции:

- иерархии абстракций (БД) [8, 9];
- агрегации (БД) [8, 9];
- обобщения (БД) [8, 9];
- нормализации [8, 9];
- семантические сети (ИИ, открытые системы и др) [10];
- сети фреймов (САПР, САД, СППР и другие) [11, 12];
- сети мультимедийных фреймов (системы дистанционного управления и мониторинга) [13];
- сети фреймов с нечеткой логикой (системы мониторинга и др.) [14].

При построении программного обеспечения применяют:

- объектно-ориентированное программирование [4] ;
- процедурное программирование [15];
- структурное программирование [16];
- модульное программирование [17];
- и другие.

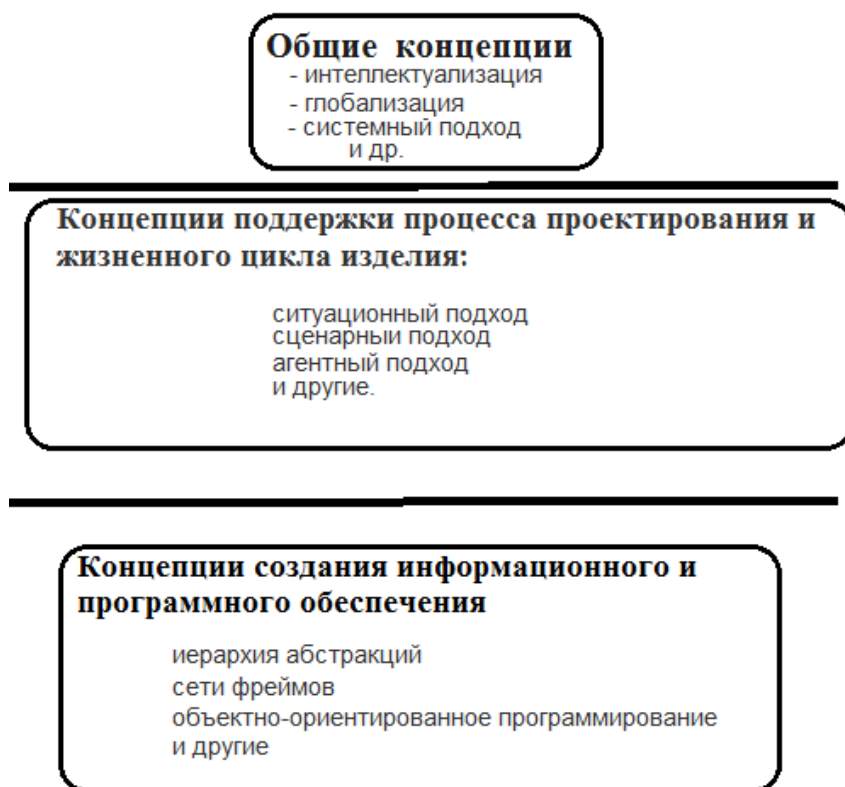


Рис. 2. Иерархическая схема концепций построения автоматизированных систем

Отметим, что описанная выше иерархия весьма условна. Например, ситуационный подход можно отнести к самому верхнему уровню иерархии т.к. он является одним из основных подходов построения ИТ систем (первые центры ситуационного управления появились в 70-х годах прошлого столетия и сегмент рынка данного класса систем сегодня активно расширяется). Однако, ситуационный подход можно отнести и к более низким уровням иерархии т.к. он активно применяется при построении автоматизированных систем (САПР, САД, систем дистанционного управления и мониторинга, систем мониторинга ситуационного типа).

С практической точки зрения для предложенной систем работает правило: чем выше уровень концепции, тем большие предметные области автоматизации он охватывает. Кроме этого проектировщик, создатель систем имеет большую свободу действий, т.к. концепции более высокого уровня не предлагают конкретных стратегий создания систем.

Например, очевидная концепция – тренд интеллектуализации автоматизированных систем и всего ИТ мира не дает технологии решения конкретной задачи. Проектировщик при создании конкретной системы в рамках концепции интеллектуализации столкнется с проблемой выбора:

- использовать интеллектуальный анализ данных;
- создавать систему интеллектуальной поддержки пользователя;
- использовать модели данных искусственного интеллекта для построения информационного обеспечения;
- применение различных механизмов логического вывода;
- использование гибридных знаний для решения задач;
- и другие.

Иерархия моделей

Анализ моделей поддержки процесса проектирования автоматизированных систем и их поддержки в течение всего ЖЦ так же позволяет их ранжировать.

На самом верхнем уровне расположим модели:

- описания процессов проектирования и поддержки систем в течение жизненного цикла (каскадная, спиралевидная и др. модели) [18];
- динамическая модель поддержки выбора ресурсов для синтеза проектируемой системы [19];
- динамическая модель описания процесса реализации сценариев функционирования автоматизированных систем [20].

К моделям более низкого уровня, которые имеют ярко выраженную инженерную ориентацию отнесем:

- фреймовая модель описания процесса проектирования изделий [12];
- модель с мультимедийными фреймами [13];
- фреймовая модель с процедурами нечеткой логики [14];
- фреймово-продукционная модель представления процессов проектирования [21];
- ситуационные модели смешанного типа [22];
- и другие.

Таким образом, совокупность моделей поддержки процесса проектирования и жизненного цикла изделия на уровне концепции образует два уровня иерархии (рис.3).

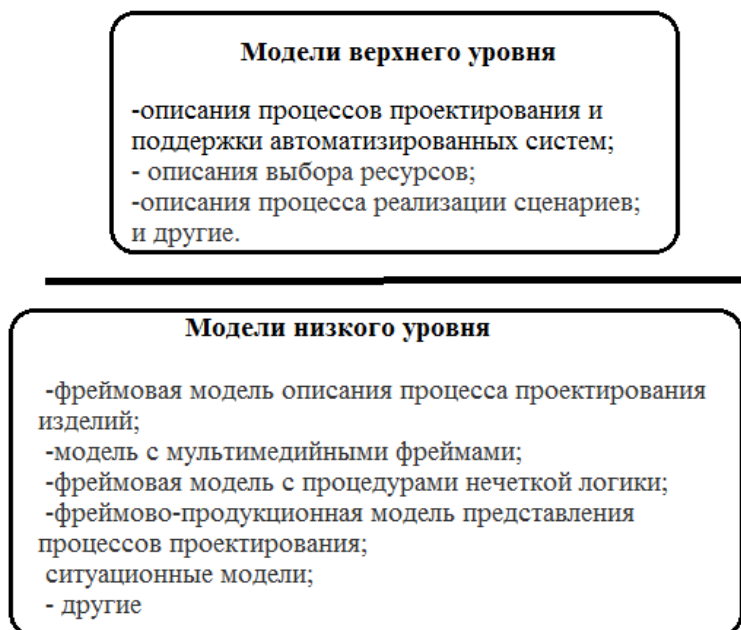


Рис. 3. Иерархия моделей описания процессов проектирования и поддержки автоматизированных систем в течение всего жизненного цикла

Отметим, что предложенная классификация концепций имеет три уровня, а классификация моделей состоит из двух уровней иерархии. Это во многом определяется традицией: при проектировании выделяют три уровня абстракций. На верхнем уровне стоят кибернетические концепции, которые описывают процесс в предметной области. Более низким уровнем являются машинно-независимые концепции и концептуальные модели, которые позволяют реализовать проект с использованием нескольких

альтернативных платформ (или инструментальных комплексов, например, СУБД). И, наконец, на самом низком уровне располагаются машинно-ориентированные концепции и модели, которые поддерживают реализацию проекта в жестких рамках конкретных инструментариев.

Библиографический список

1. Клименко И. Теория систем и системный анализ. Учебное пособие. Кнорус, 2021. 264 с.
2. Юрчик П.В., Голубкова В.Б. Применение CALS технологий на предприятии. Учебное пособие/ Лань, 2020 г. – 92 с.
3. Гуляев Ю.В., Журавлев Е.Е., Олейников А.Я. Методология стандартизации для обеспечения интероперабельности информационных систем широкого класса. Радиоэлектроника, N3, 2012, С 1-12.
4. Буч Г. Объектно-ориентированный анализ и проектирование с примерами приложений 3-е издание, -М.: ООО "И. Д. Вильямс", 2008. - 720 с.: ил.
5. Поспелов Д.А. Ситуационное управление: теория и практика. М.: Наука, 1986. — 288 стр.
6. Мисевич П.В. Логистический подход к вопросам проектирования систем объектов и формирования сценариев работы программного обеспечения// Системы управления и информационные технологии. 2007. №4.1(30). - С. 169-174.
7. Рассел С. Искусственный интеллект: современный подход (AIMA) / Стюарт Рассел, Питер Норвиг. - М.: Вильямс, 2007. - 1424 с.
8. Тиори Т. Проектирование структур баз данных. В 2 кн./ Т. Тиори, Д. Фрай - М.: Мир, 1985. Кн. 1. - 287 с. Кн. 2. - 320 с.
9. Паклин Н.Б., Орешков В.И. Бизнес-аналитика: от данных к знаниям. Учебное пособие. — 2-е изд., испр. — СПб.: Питер, 2013. — 704 с.
10. Осипов Г.С. Методы искусственного интеллекта / Осипов Г. С. - Москва: ФИЗМАТЛИТ, 2011. - 296 с.
11. Минский М. Фреймы для представления знаний / М. Минский. - М: Энергия, 1979. - 151 с.
12. Семёнов В.В. Принципы формирования и фрагменты базы знаний теории управления, в сборнике "Общее математическое обеспечение систем автоматизированного проектирования", Москва, МАИ, 1981.
13. Белов Д.А. Проблемно-ориентированная автоматизированная система мониторинга движения железнодорожного состава / Д.А. Белов, П.В. Мисевич, В.П. Хранилов // Автоматизация в промышленности. 2009. №2. С.49-51.
14. Ермилов А. Э., Мисевич П. В. Применение фреймовой модели и нечёткой логики в основе построения инструментариев автоматизированных систем мониторинга. // Труды НГТУ им. Р. Е. Алексеева, 2015, №1(108), Нижегород. гос. техн. унт. -Н. Новгород, С.71-76.
15. Олейников Б. Программирование. процедурное программирование издательство: Сибирский федеральный университет 2016. – 92 с.
16. Павловская Т., Щупак Ю. С/С++. Структурное и объектно-ориентированное программирование: Практикум СПб.: Питер, 2011. — 352 с.
17. Титовская Н.В., Титовский С.Н Модульное программирование: учеб. пособие / Н.В.Титовская; С.Н.Титовский, Краснояр. гос. аграр. ун-т. -Красноярск
18. Сольнищев Р.И. Автоматизация проектирования систем автоматического управления / Р.И. Сольнищев - М., Высшая школа, 1991. - 335с.
19. Хранилов В.П. Нечеткая динамическая модель интерактивного распределения вычислительных ресурсов / В.П. Хранилов, Д.В. Прохоров // Системы управления и информационные технологии. 2006. № 4.1(26). С. 189-12020

20. Мисевич П.В. Динамическая модель функционирования автоматизированной системы // Системы управления и информационные технологии, 2008, 3.1(33). - С. 175-179.
21. Басалин П.Д. Построение САПР на базе фреймово - производственной модели / П.Д. Басалин, П.В. Мисевич // Системы обработки информации и управления: межвуз. сб. научн. тр. / Н. Новгород: изд - во НГТУ, 1995. - С. 106 - 112.
22. Полевая О. М. Математическое обеспечение синтеза формулировок стратегических целей и задач в информационной системе поддержки процессов стратегического управления // Информационные системы и технологии, №3, 2016, с.81-91.

P.V. Misevich, A.E. Ermilov, E.N.Pankratova, V.P. Khranilov

THE HIERARCHY OF CONCEPTS AND MODELS FOR DESIGNING AUTOMATED SYSTEMS

Nizhny Novgorod state technical university n.a. R.E. Alekseev

The paper presents a hierarchical system of concepts for building automated systems. All concepts are divided into three levels of hierarchy. At the top of the classification system there are concepts for describing global trends in the world information technology. At the middle level there are concepts for supporting the design process. Concepts for building specific types of automated systems or software, DB and knowledgebase are situated at the bottom of the hierarchical system.

Also in the article it is proposed to use a classification system of models for supporting the design process. There are two levels of hierarchy.

The universal declarative models for describing the design process are situated at the first level. These models are relevant to every design process. Project-oriented models for supporting design operations and procedures are situated at the second level.

Keywords: automated system, model, design, life cycle, frame network, fuzzy logic, hybrid knowledge, frame-production model, frames with multimedia images, control systems, monitorins system.

ИДЕНТИФИКАЦИЯ ПЛОТНОСТЕЙ РАСПРЕДЕЛЕНИЯ СЛУЧАЙНЫХ ВЕЛИЧИН. СТЕПЕННОЙ ФУНКЦИОНАЛЬНЫЙ БАЗИС И РЕЗУЛЬТАТЫ ЕГО ТЕСТИРОВАНИЯ

Нижегородский государственный технический университет им. Р.Е. Алексеева

Для идентификации полимодальных распределений случайных величин наиболее эффективны методы, основанные на решении обратных задач математической физики (приближенное решение уравнений Фредгольма первого рода). При этом используется традиционный тригонометрический базис представления искомым плотностей распределения. Авторами статьи показана также целесообразность и область применения полиномиального базиса Лежандра.

В настоящей публикации рассмотрены возможности неортогонального степенного базиса. Показано, что базис можно использовать для экспресс-идентификации и восстановления распределений, подобных равномерному. Этот вариант требует применения методов регуляризации для эквивалентных систем уравнений.

Ключевые слова: идентификация распределений, полимодальные плотности вероятности, уравнение Фредгольма первого рода, степенной функциональный базис.

Введение

В публикации [1] приведены результаты идентификации распределений прикладного характера (биомедицина), основанные на двух функциональных ортогональных базисах: первый – тригонометрический, второй – Лежандра. Базисы имеют высокое разрешение и могут эффективно идентифицировать полимодальные распределения по выборкам как малого, так и большого объема [2-4].

Вместе с тем следует отметить, что тригонометрический базис, обладая высоким быстродействием, имеет недостаток, выражающийся в том, что для распределений с равномерной плотностью или уплощенного вида дает трудно устранимую колебательную составляющую.

С восстановлением таких плотностей хорошо справляется метод на основе базиса Лежандра. Однако скорость вычислений в общем случае у него значительно ниже (на порядок и более) и составляет несколько десятков секунд.

Возникает задача исследовать компромиссный вариант, сочетающий плюсы обоих базисов. В качестве одного из вариантов предлагается исследовать идентификационные возможности простого степенного полинома. Указанный базис не является ортогональным. Поэтому восстановление коэффициентов полинома, аппроксимирующего искомую плотность распределения, становится потенциально неустойчивой, некорректной задачей. Для минимизации функционалов среднего риска в этом случае необходимо применять методы регуляризации эквивалентных систем уравнений.

Для решения задачи предложено использовать регуляризирующий ЕС-алгоритм, рассмотренный и модифицированный в работах [5,6].

Ниже приведены результаты тестирования метода со степенным базисом и встроенным в программу идентификации ЕС-алгоритмом регуляризации. Для верификации возможностей третьей системы функций выполнена идентификация равномерного распределения и нормально распределенных случайных величин с различной дисперсией.

Результаты тестирования и особенности простого степенного базиса

На рис.1 (а, б) показаны результаты идентификации одномодальных (с различными дисперсиями) гауссовых плотностей, сгенерированных в пакете Матлаб. В целом – результат удовлетворительный, форма распределений восстановлена с малыми искажениями, время расчетов – примерно 1 с. Энтропийный коэффициент $k \approx 2,045$. Теоретическое значение – 2,066.

Еще меньшее время и весьма точное восстановление показано для равномерных распределений. В этом случае результаты идентификации со степенным базисом полностью совпали с результатом базиса Лежандра (рис.1 в, г).

Предварительный вывод тестовой идентификации: степенной базис имеет высокое быстродействие и эффективен при восстановлении одномодальных плотностей вероятности. Регуляризация обеспечивает устойчивость.

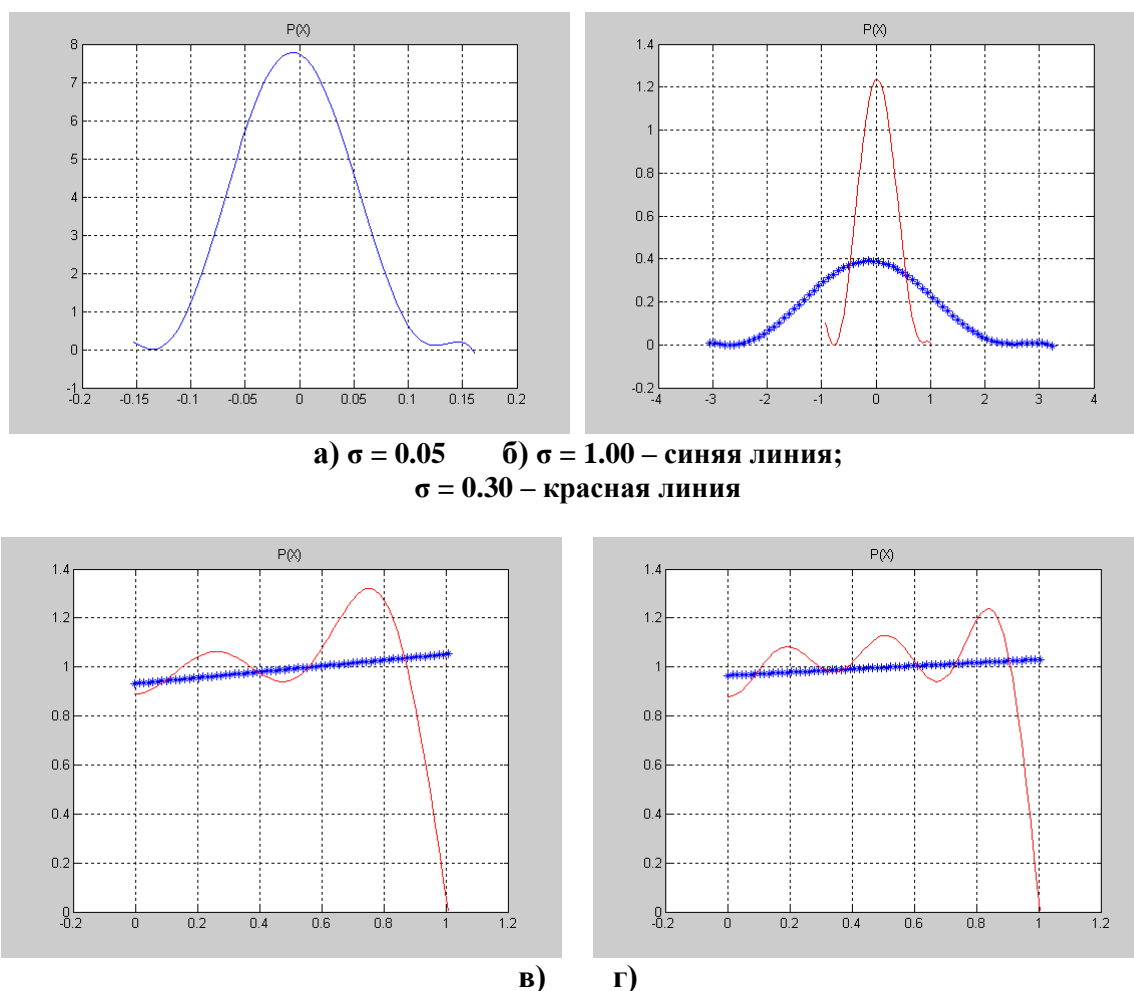


Рис. 1. Идентифицированные плотности распределения тестовых случайных величин, распределенных по нормальному закону (объем выборки $L = 1200$) с нулевым математическим ожиданием (а, б). Функциональный базис – степенной полином; (в, г) - идентифицированные плотности равномерного распределения на отрезке $[0,1]$ (объем выборки соответственно 1200 и 4800); жирная линия – базисные функции – полиномы Лежандра и степенной базис, тонкая – тригонометрический базис.

Исследованы два сложных варианта полимодальных распределений с уединенными двумя и тремя модами.

Вариант исследования распределения с двумя разнесенными модами показан на рис. 2. В этом примере алгоритмически объединены нормальные распределения с центрами группирования в точках $x = 0$ и $x = 12$. Встроенный в программу идентификации

регуляризирующий модуль позволяет восстановить обе моды, хотя и не в полной мере. Вычислительный алгоритм работает в «жестких» условиях больших значений чисел обусловленности и практически абсолютной вырожденности, решаемых СЛАУ (рис. 2 в, г).

Приведем ряд значения чисел обусловленности решаемых систем: 11.721; 1451.7; 75989; 1.4709e+007; 2.3154e+008; 1.5716e+010 . . . 1.7857e+012.

Очевидно, что достаточно успешно решена существенно плохо обусловленная задача. Найденный оптимальный параметр регуляризации $\alpha_{\text{опт}} = 4.40\text{e-}004$, время идентификации составляет примерно 3 с.

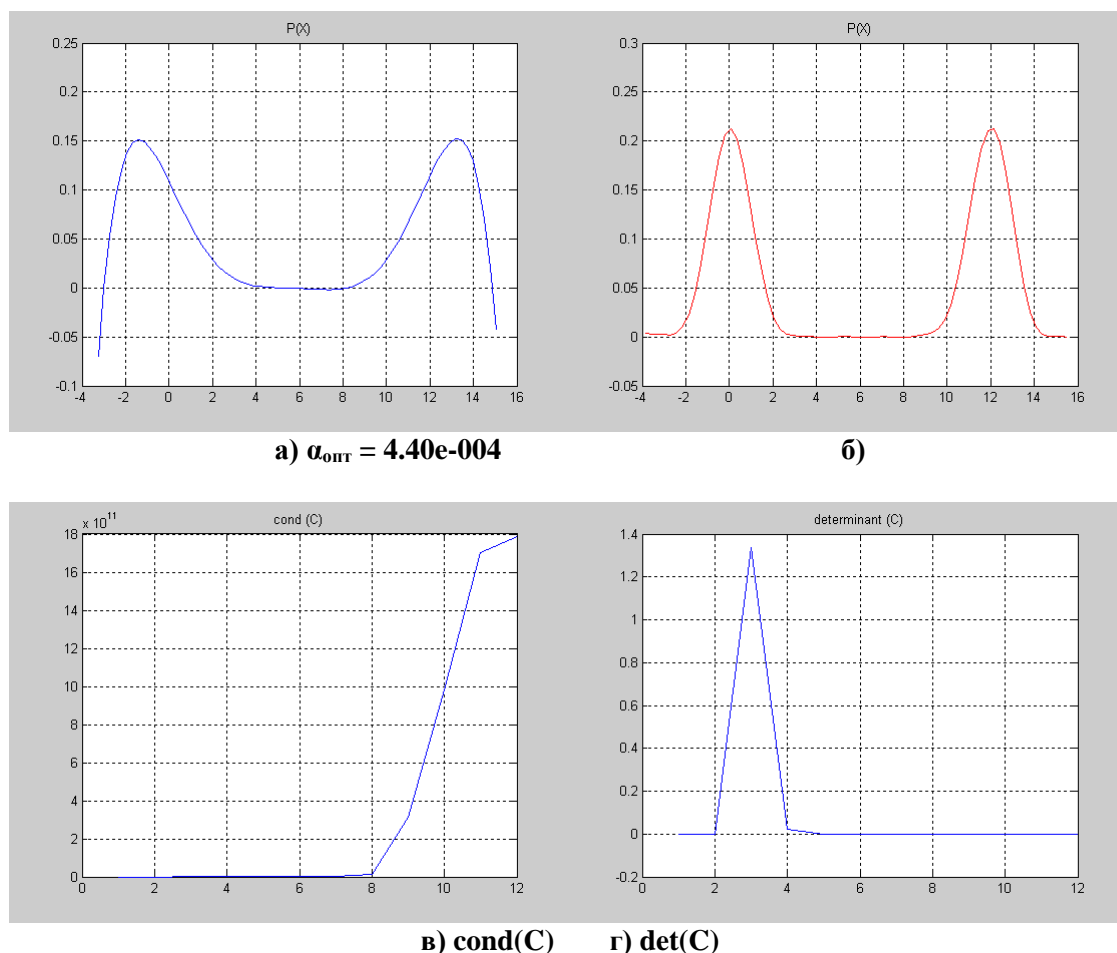
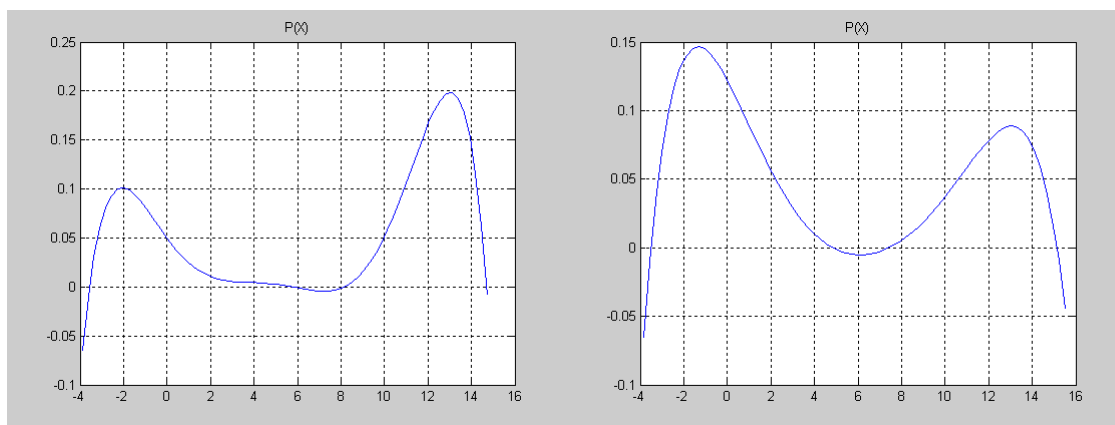


Рис. 2. Идентифицированные плотности распределения бимодальной случайной величины с модами, распределенных по нормальному закону ($L = 2400$).

При этом, на рисунке: а) функциональный базис – степенной полином; б) функциональный базис – тригонометрический. Графики (в, г) показывают число обусловленности и детерминант решаемых в процессе идентификации СЛАУ. По оси абсцисс – порядок системы уравнений N . $N_{\text{MAX}} = 12$.

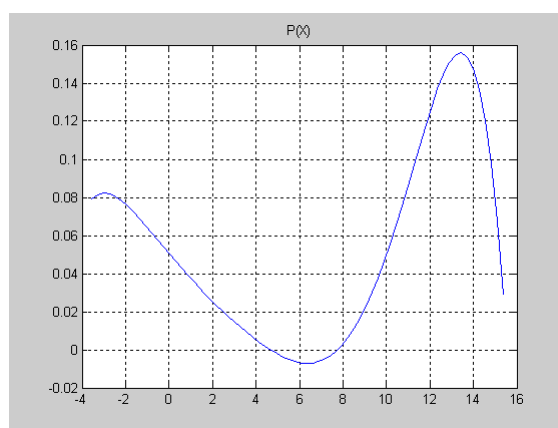
Рис. 3 демонстрирует результаты решения задачи без регуляризации и при не оптимальных множителях Лагранжа α .

Искажение формы гауссианов – следствие плохой обусловленности решаемых СЛАУ на этапе безусловной минимизации функционала среднего риска. Это свойство неортогонального базиса подтверждают аналогичные результаты идентификации трехмодального тестового распределения – координаты центров распределения $x = 0$, $x = 12$, $x = 24$ (рис. 4).



а) $\alpha = 0$

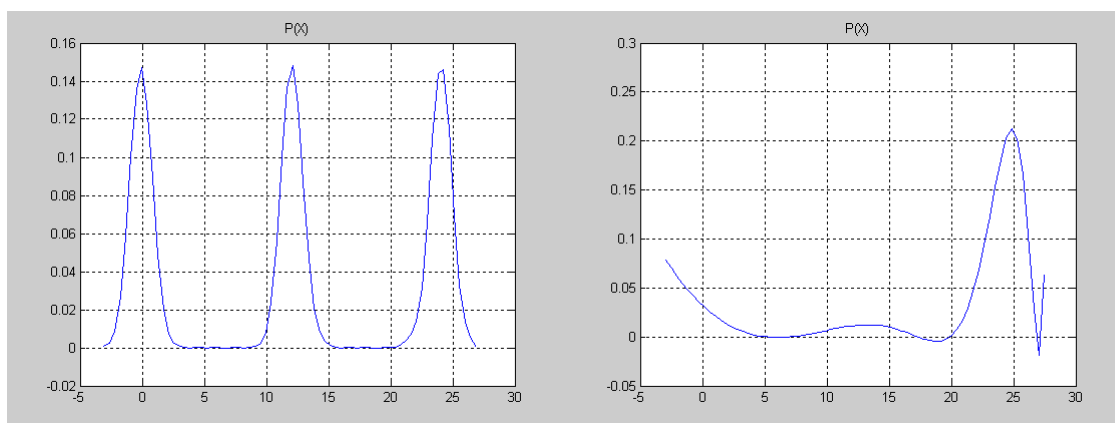
б) $\alpha = 1.e-09$



в) $\alpha = 1.e-06$

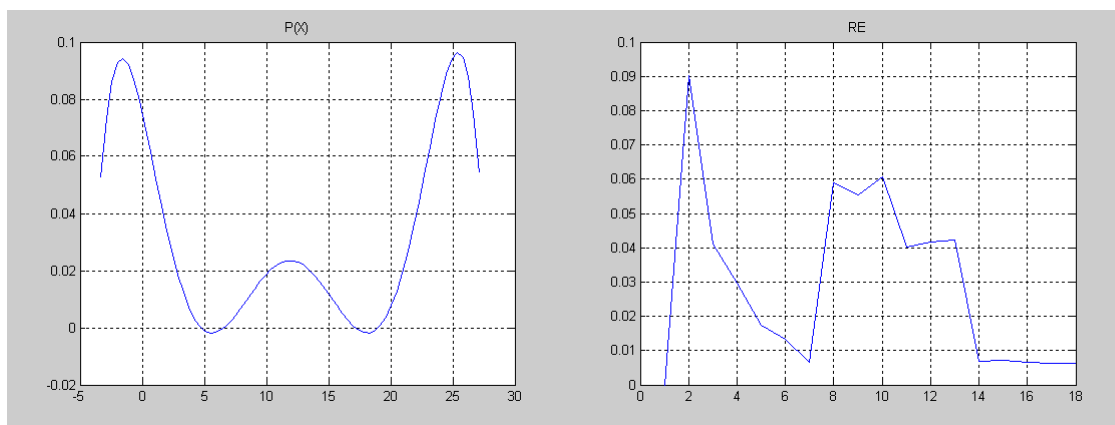
Рис. 3. Восстановленные плотности распределения бимодальной случайной величины без регуляризации – (а) и с тестовыми (неоптимальными) параметрами регуляризации α – (б) и (в). Функциональный базис – степенной полином.

Время идентификации 7 оптимальных коэффициентов разложения искомой трехмодальной плотности (соответственно и степени полинома) составило 5.7 с. С учетом сложности теста – быстрый спад значений сингулярного спектра решаемых СЛАУ до нуля и значения числа обусловленности $\text{cond}(C) \approx 1. e+14$ – результат идентификации следует признать удовлетворительным. Выявлены все три моды и их «координаты» точно обозначены. «Дотянуть» степень восстановления до оптимума корректности и сохранить быстрдействие предположительно можно, если применить полиномы Бернулли.



а)

б)



в) $\alpha = 4.883e-004$ $N_{opt} = 7$ г)

Рис. 4. Идентификация плотности распределения трехмодальной случайной величины – тригонометрический базис – (а). Результаты восстановления плотности без регуляризации (б) и с регуляризацией ЕС-алгоритмом (в) (функциональный базис – степенной полином). График значений минимизируемого функционала (г).

Выводы

При минимизации функционала среднего риска для неортогонального степенного базиса необходимо использовать методы регуляризации эквивалентных систем уравнений.

Степенной базис можно применять для распределений с уединенными модами и числами обусловленности решаемых систем до 10^8 , а в общем случае, – для экспресс-идентификации и восстановления распределений, подобных равномерному.

В качестве неортогональных степенных базисов также целесообразно исследовать на устойчивость, степень разрешения и быстродействия полиномы Эйлера и Бернулли.

Исследование выполнено при финансовой поддержке Российского фонда фундаментальных исследований в рамках научного проекта № 19-07-00926_a.

Библиографический список

1. Куликов В.Б., Куликов А.Б., Хранилов В.П. Идентификация и верификация законов распределения биомедицинских показателей на основе ортогонального базиса Лежандра // Труды Нижегородского государственного технического университета им. Р.Е. Алексева. 2020. Т. 131. № 4. С. 7-15.
2. Куликов В.Б. Восстановление полимодальных плотностей вероятности по экспериментальным данным в структурах со стохастическими свойствами // Вестник Нижегородского университета им. Н.И. Лобачевского. 2014. № 1(1). С. 248-256.
3. Тихонов А.Н., Арсенин В.Я. Методы решения некорректных задач. М.: Наука, 1986. 288 с.
4. V. Kulikov. The Identification of the Distribution Density in the Realization of Stochastic Processes by the Regularization Method // Appl. Mathem. Sciences, vol. 9, № 137, 2015, pp. 6827-6834.
5. Куликов В.Б., Куликов А.Б. Идентификация плотностей распределения характеристик стохастических структур. Алгоритмы регуляризации // Математические методы в технике и технологиях (ММТТ-32): Сб. тр. междунар. науч. конф.: в 12 т. Т. 3 / под общ. ред. А. А.Большакова. - СПб.: Изд-во Политехн. ун-та, 2019. С. 94-98.
6. Kulikov V., Kulikov A. (2020) Regularization Methods for the Stable Identification of Probabilistic Characteristics of Stochastic Structures. In: Kravets A.,

Bolshakov A., Shcherbakov M. (eds) Cyber-Physical Systems: Advances in Design & Modelling. Studies in Systems, Decision and Control, vol 259. pp. 179-191. Springer, Cham.

V. B. Kulikov, A. B. Kulikov, V. P. Khranilov

IDENTIFICATION OF RANDOM VARIABLE DISTRIBUTION DENSITIES. THE POLYNOMIAL FUNCTIONAL BASIS AND RESULTS OF ITS TESTING

Nizhny Novgorod state technical University n. a. R. E. Alekseev

Purpose: To identify polymodal distributions of random variables, the possibilities of a non-orthogonal power basis are considered.

Methodology/approach: The traditional identification method, based on two functional orthogonal bases: the trigonometric and Legendre basis, is supplemented by a compromise option that combines the advantages of both bases. This basis is not orthogonal, so to minimize the mean risk functionals in this case, it is necessary to use regularization methods of equivalent systems of equations.

Findings: For a stable solution of the problem, the authors propose a regularizing RRA-algorithm.

To verify the capabilities of the system of power functions, the identification of a uniform distribution and normally distributed random variables with different variance is performed.

Complex variants of polymodal distributions with isolated two and three modes are also studied.

Value: The power basis can be used for distributions with solitary modes and conditionality numbers of the solved systems up to 10^8 , and in the general case, for express identification and reconstruction of distributions similar to the uniform one.

Keywords: identification of distributions, polymodal probability densities, Fredholm equation of the first kind, polynomial functional basis.

УДК 621.039:681.5

И.Ю. Лепехин, П.А. Голубев, В.Г. Промыслов, К.В. Семенков

ПРОБЛЕМЫ ЭКСПЕРИМЕНТАЛЬНОГО ПОДТВЕРЖДЕНИЯ ВРЕМЕННЫХ ХАРАКТЕРИСТИК ЦИФРОВЫХ СИСТЕМ УПРАВЛЕНИЯ

*Институт проблем управления им. В.А. Трапезникова РАН, Россия, г. Москва
Филиал «ФГУП РФЯЦ-ВНИИЭФ» «НИИИС им. Ю.Е. Седакова», г. Нижний Новгород*

В работе рассмотрена проблема практического подтверждения временных характеристик цифровой системы управления. Показано, что задержка обработки и передачи информации в общем случае имеет функцию распределения отличную от нормальной, следовательно, доверительные интервалы, рассчитанные в предположении о нормальном распределении, будут не верны.

Предполагается для корректного подтверждения учитывать особенности распределения или применять нестатистические методы оценки временных параметров. Например, использовать метод Network Calculus.

Ключевые слова: испытания, система управления, атомные станции, временные характеристики.

Введение

Современная система управления (СУ) в промышленных системах в большинстве случаев реализуется в виде распределенной по функциям и средствам вычислительной системы с компонентами, представляющими собой совокупность аппаратных и программных средств, которые обеспечивают сбор, накопление, асинхронную обработку, представление и передачу информации. Компоненты СУ могут быть распределены как пространственно, так и функционально.

Для подтверждения работоспособности СУ разработаны методы статической и динамической валидации систем; например, для СУ в АЭС применяется методика МАГАТЭ [IAEA, 2017]. Частью испытаний системы является подтверждение ее временных характеристик с точки зрения, как быстродействия (времени обработки) данных в компонентах системы, так и времени (задержки) передачи данных между компонентами.

В основном, используются статистические методы подтверждения временных характеристик СУ в ходе их валидации. При этом распространено предположение о близком к нормальному распределению измеряемых характеристик [IAEA, 2017], что позволяет надежно оценить временные параметры СУ при достаточном объеме выборки. Данное предположение в основном верно для сигналов в системе, которые имеют физическую природу, но, как мы покажем, не выполняются для величин, описывающих саму цифровую СУ. Например, оно не выполняется для величины задержки и времени обработки сигнала.

Нами исследовались основные статистические свойства временных характеристик типовой СУ для АЭС на натурном макете системы верхнего блочного уровня разработки ИПУ РАН [Бывайков, 2006].

Типовая система управления и модель данных

Типовая системы управления большинства промышленных объектов содержит, по крайней мере, четыре логических уровня:

- уровень человеко-машинного интерфейса;
- уровень алгоритмов управления и хранения данных;
- уровень низовых контролеров сбора и передачи информации;
- уровень исполнительных механизмов и датчиков.

Данные логические уровни могут объединяться для небольших систем, но выполняют свою определенную функцию.

В типовой структуре нами исследовались только полностью цифровые компоненты структуры, приведенные на схеме (Рис. 2):

- низовые контролеры - G уровень (1);
- серверы первичной обработки и хранения данных - S уровень (2);
- компоненты, обеспечивающие человеко-машинный интерфейс - Z уровень (3).

Алгоритмы обработки информации компонентов G, S, Z циклически.

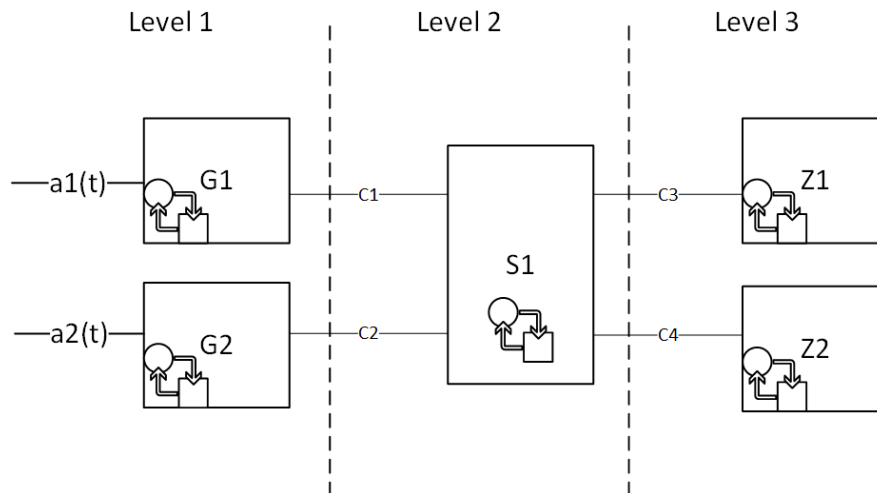


Рис. 2. Структурная схема исследуемой системы управления.

В соответствии с практическими условиями эксплуатации СУ (Рис. 2) на АЭС, между уровнями АСУ ТП и компонентами в них существует циклический поток данных, который представляет собой значения сигналов о состоянии объекта управления и самой СУ, передаваемые от шлюза на уровне присоединения к серверу и от сервера к рабочей станции в определенный интервал времени.

Циклические данные представлены в виде:

$$A = f(L_c, N_c, D_c)$$
$$D = T_S + T_E,$$

где L_c - размер циклических данных, которые содержат заголовок кадра, поле адреса, поле данных, поле контроля четности и так далее;

N_c - количество циклических данных, поступающих в единицу времени;

D_c - это полное время задержки при передаче пакета, представляющее сумму задержки в сети Ethernet T_E и времени обработки на вычислительном ресурсе T_S .

В работе исследовались только свойства T_S, T_E , напрямую влияющие на временные характеристики системы.

Статистические характеристики задержек

Мы исследовали временные характеристики реальной системы [Бывайков, 2006] с архитектурой, представленной выше (Рис. 2). В работе данные приведены только для одного компонента так как для остальных компонентов параметры аналогичны.

Характеристики задержки при циклической обработке данных

Рассмотрим оценку максимального времени обработки данных в компоненте с циклическим алгоритмом функционирования.

Измерения проводились для компонент уровня Z (Рис. 2). Объем обрабатываемых циклических данных сравнительно стабилен в нормальных режимах работы и имеет среднюю скорость. Однако при некоторых особых (срабатывание защит и блокировок) или переходных (смена режимов) процессах объем данных и алгоритм (скорость) обработки могут значительно меняться.

На рисунке 2 приведена временная диаграмма циклов в нормальном режиме.

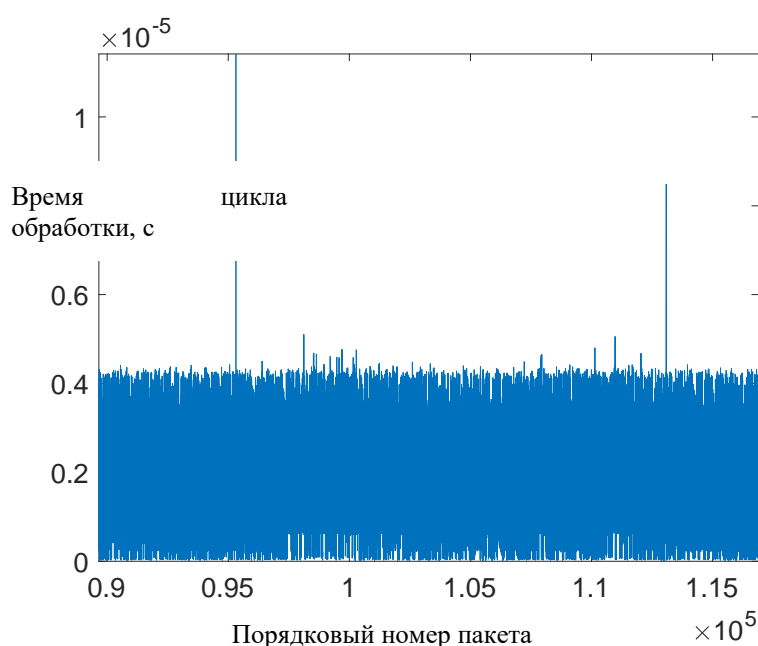


Рис. 3. Пример экспериментальной зависимости времени цикла обработки от номера пакета.

Эмпирическое распределение времени цикла T_S (Рис. 4) очевидно отличается от нормального или пуассоновского и является многомодовым. Каждая из мод соответствует характерному циклу обработки для определенного вида данных.

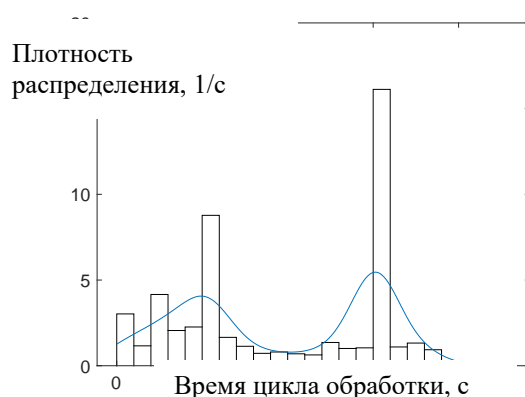


Рис. 4. Эмпирическое распределение плотности вероятности времени цикла компонента Z
Характеристики сетевой задержки при циклической обработке данных

Рассмотрим параметры сетевой задержки T_E передачи данных между компонентами СУ. Для примера приведены данные прохождения сетевых пакетов между компонентами уровня S и уровнем Z. Все компоненты в рассматриваемом примере обмениваются данными по протоколу TCP/IP. Характеристики для процесса передачи данных между компонентами других уровней системы имеют сходный характер.

В эксперименте с помощью утилиты tcpdump из состава ОС измерялось время двойного прохождения TCP-пакета (round-trip time, RTT) — время, прошедшее с момента отправки пакета компонентом S до момента получения подтверждения (ACK) от компонента уровня Z. Измеренные значения RTT приведены на Рис. 5.

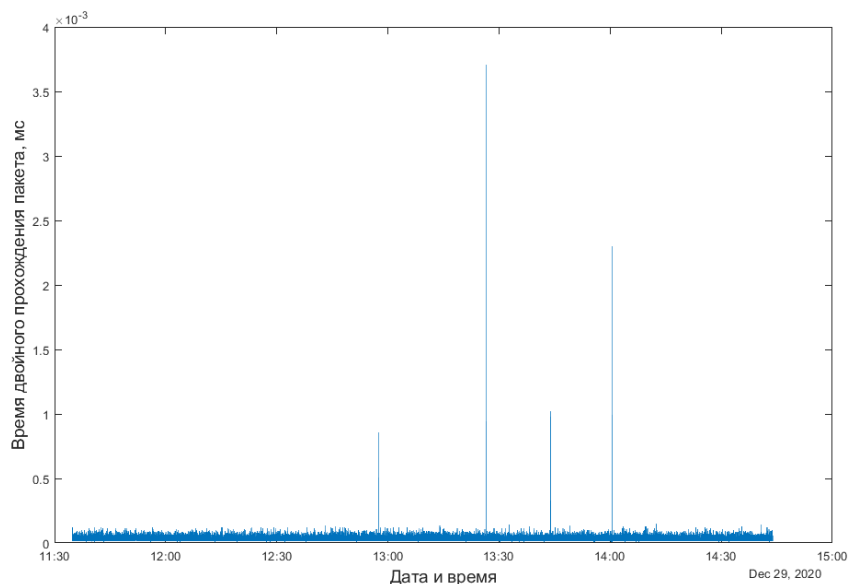


Рис. 5. Время двойного прохождения пакета между компонентами S и Z в реальной СУ.

Типичное время двойного прохождения пакета составляет десятки микросекунд. В то же время, распределение RTT существенно отличается от нормального распределения, характерного для физических процессов, или пуассоновского распределения, широко используемого в теории обслуживания (рис. 5).

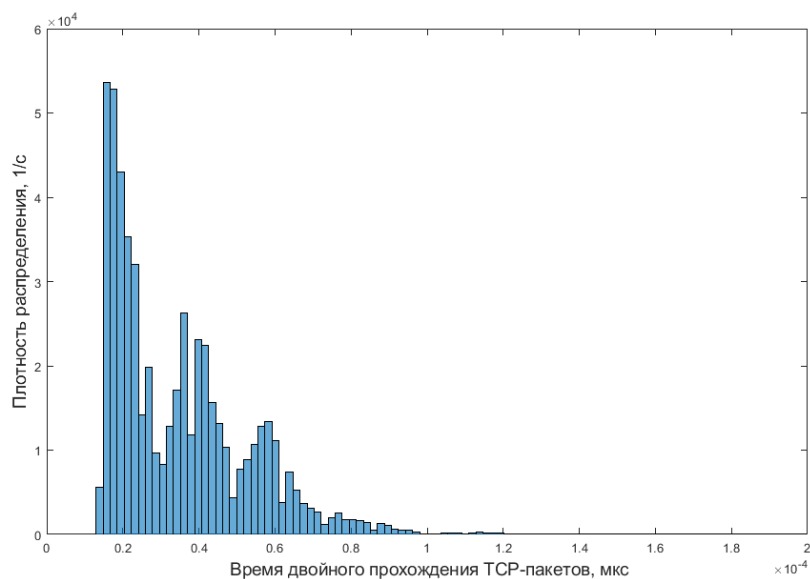


Рис. 6. Эмпирическое распределение плотности вероятности для времени двойного прохождения TCP-пакета.

На распределении (рис. 5) четко выделяются три характерных времени, распределение очевидно отличается от нормального или пуассоновского и является многомодовым.

Обсуждение результатов

Нами проведена дополнительная проверка того, что полученные в результате измерений распределения (рис. 3, рис. 5) имеют тяжелый хвост. Проверка проводилась с помощью алгоритма [Bunjeski, 2015], который у авторов показал лучшие результаты, чем тесты, основанные на критерии Колмогорова-Смирнова.

Можно заключить, что исследования на реальной системе показали, что распределения времени задержек как в сетевых компонентах, так и в компонентах, обрабатывающих информацию, отличаются от нормального, имеют частично многомодовый характер и относятся к типу распределений с тяжелыми хвостами.

Заключение

В требования по производительности СУ часто включаются:

- ограничения на время обработки в отдельных компонентах СУ,
- ограничения на время передачи информации между компонентами СУ.

Ограничения могут накладываться как на средние, так и на максимальные (предельные) значения величин и выражаться либо в виде статистических ограничений, когда вместе со величиной задаётся ее доверительный интервал, либо в виде ограничений на абсолютные значения [ГОСТ 28195-89].

Оценивание случайной величины по выборке — классическая задача математической статистики, она давно и хорошо разработана и изложена в литературе [см., например, [Крамер, 1975]]. Однако интерпретация полученных оценок, распространяющая опыт работы с «обычными» измерениями физических величин на оценку временных характеристик цифровых систем управления, может привести к неверным выводам.

Сформулируем основные проблемы.

Физическая природа измеряемой величины (времени) накладывает ограничения на вид ее функции распределения: как минимум, она является ограниченной слева.

Если в технических требованиях задается максимальное абсолютное значение, например, «время прохождения сигнала между компонентами СУ не должно превышать некоторой величины», то такая форма условия подразумевает, что случайная величина имеет функцию распределения, ограниченную также и справа, т.е. ограничения означают, что функция распределения изначально не является в строгом смысле функцией распределения нормальной случайной величины.

Процедура подтверждения требований в основном базируется на значении выборочного среднего и выборочной дисперсии [IAEA, 2017]. При условии, что случайная величина имеет конечное математическое ожидание и дисперсию, выборочное среднее является несмещённой состоятельной оценкой теоретического среднего и не зависит от вида распределения. Его известным недостатком [Smith, 2015] является низкая робастность в случае наличия посторонних выбросов в выборке. Выборочная дисперсия, как смещенная, так и несмещенная, также является состоятельной оценкой теоретической дисперсии величины.

Если бы случайная величина имела нормальное распределение, то имея выборочное среднее и дисперсию, несложно точно оценить доверительный интервал для подтверждаемого параметра. Однако, как показано ранее, функции распределения задержки в СУ в общем случае не является нормальной поэтому интуитивно, перенося оценки доверительных интервалов для нормально распределенной величины, на случай временных задержек в системах управления, получают ложные результаты.

В общем случае для оценки вероятности превышения случайной величиной некоторого значения можно воспользоваться неравенством Чебышева, но оно дает очень грубую оценку. Поэтому при оценке временных параметров необходимо исследовать функцию распределения и при необходимости использовать соответствующие статистические методы оценки для данного класса распределений. Также можно применять детерминированные методы оценки временных параметров, например, использовать метод Network Calculus [Le Boudec 2019]. Network calculus предлагает удобную методику расчета времени прохождения сигналов (задержек) в цифровых потоковых сетях. Основные понятия, необходимые для решения этой задачи, следующие. Рассмотрим систему без потерь как черный ящик. Входящий и выходящий потоки моделируются кумулятивными функциями, т.е. суммарным количеством данных, наблюдаемых на входе и на выходе за определенный интервал времени. Время прохождения сигнала (задержка) определяется как горизонтальное отклонение между этими функциями. Максимальное значение времени прохождения сигнала может быть вычислено с помощью соответствующих ограничений, а не фактических кумулятивных функций. Для этой цели используется верхняя граница или огибающая входящего потока и нижняя граница или минимальная функция обслуживания. При этом максимальное время прохождения сигнала равно горизонтальному отклонению между этими ограничениями. Данный метод успешно применялся для расчета временных характеристик современных систем управления [Масолкин, 2010], [Baybulatov, 2020].

Библиографический список

1. Крамер Г. Математические методы статистики. – М.: Мир, 1975.
2. Бывайков М.Е., Жарко Е.Ф., Менгазетдинов Н.Э., Полетыкин А.Г., Прангишвили И.В., Промыслов В.Г. Опыт проектирования и внедрения системы верхнего блочного уровня АСУ ТП АЭС // Автоматика и телемеханика. 2006. Т. 5. с. 65-79.
3. Масолкин С.И., Промыслов В.Г. Расчет некоторых параметров промышленной вычислительной сети объектов повышенного риска эксплуатации на примере АСУТП АЭС // Проблемы управления. 2010. с 53.
4. ГОСТ 28195-89 Оценка качества программных средств. Общие положения. 1989
5. Ahmad Z. and Durad M. H., "Development of SCADA Simulator using Omnet++," 2019 16th International Bhurban Conference on Applied Sciences and Technology (IBCAST), Islamabad, Pakistan, 2019, pp. 676-680, doi: 10.1109/IBCAST.2019.8667158.
6. Baybulatov A. and Promyslov V., "On a Deterministic Approach to Solving Industrial Control System Problems," 2020 International Russian Automation Conference (RusAutoCon), Sochi, Russia, 2020, pp. 115-120, doi: 10.1109/RusAutoCon49822.2020.9208149.
7. Burnecki K, Wylomanska A, Chechkin A (2015) Discriminating between Light- and Heavy-Tailed Distributions with Limit Theorem. PLoS ONE 10(12): e0145604. doi:10.1371/journal.pone.0145604 2015
8. IAEA TECDOC No. 1830 On-line Monitoring of Instrumentation in Research Reactors 978-92-0-108517-7 2017
9. Le Boudec, J-Y., and Thiran, P. (2019). Network Calculus: A Theory of Deterministic Queuing Systems for the Internet, Springer Verlag, Online Version of the Book, LNCS 2019.
10. Promyslov V. Using the Method of "Network Calculus" to Simulate the Computerized Control System with Non-Uniform Data Flow. IFAC Proceedings Volumes (IFAC-PapersOnline). 14. 645-648. 10.3182/20120523-3-RO-2023.00273, 2012.

11. Smith M. Statistical Analysis Handbook a Comprehensive Handbook of Statistical Concepts, Techniques and Software Tools Copyright © 2015-2018 All Rights reserved. 2018 Edition. Issue version: 2018-1
12. Golshani G., Taylor G., and Pisica I., "Simulation of power system substation communications architecture based on IEC 61850 standard," in 2014 49th International Universities Power Engineering Conference (UPEC), 2014, pp. 1-6.

Golubev P.A., Lepekhin I.Iu., Promyslov V.G., Semenov K.V.

PROBLEMS OF EXPERIMENTAL VALIDATION TIMING CHARACTERISTICS OF DIGITAL CONTROL SYSTEMS.

*Branch of the RFNC-VNIIEF "NIIS n.a. Yu.Ye. Sedakov", Nizhniy Novgorod
V. A. Trapeznikov Institute of Control Sciences of Russian Academy of Sciences*

The paper considers the problem of practical confirmation of the time characteristics of digital control systems. It is shown that the important time characteristic a delay in the general case has a distribution function different from the normal one. Therefore, the confidence intervals calculated in the normal distribution proposal will not be correct.

For correct confirmation, it is proposed to take into account the actual distribution features or use non-statistical methods for estimating time parameters. For example, use the Network Calculus method.

Keywords: testing, control system, nuclear stations, validation, delay, time characteristic.

САМОДИАГНОСТИКА И ВОССТАНОВЛЕНИЕ ПРОМЕЖУТОЧНЫХ РЕЗУЛЬТАТОВ ВЫЧИСЛИТЕЛЯ НА ОСНОВЕ КОММУТАТОРА SPACEWIRE 1907BM056

Филиал ФГУП «РФЯЦ-ВНИИЭФ» «НИИИС им. Ю.Е. Седакова», г. Нижний Новгород

Показан трехканальный вычислитель с «холодным» резервированием на основе СБИС многопортового коммутатора SpaceWire 1907BM056. Разработаны алгоритмы поддержания правильности функционирования вычислителя за счет проведения самодиагностики, управления резервированием и восстановления из специализированного ОЗУ промежуточных результатов вычислений. Подтверждено взаимодействие модулей вычислителя по интерфейсу SpaceWire при организации доступа к памяти по протоколу RMAP.

Ключевые слова: самодиагностика, восстановление, SpaceWire, RMAP, 1907BM056

Объект и предмет исследований

Объектом исследований является трехканальный вычислитель с «холодным» резервированием, разработанный и изготовленный в филиале РФЯЦ-ВНИИЭФ «НИИИС им. Ю.Е. Седакова» (рисунок 1).



Рис. 1. Трехканальный вычислитель с «холодным» резервированием

Предметом исследований является поддержание правильности функционирования вычислителя в процессе эксплуатации, в том числе, обнаружение сбоев или отказов, управление резервированием и восстановление вычислительного процесса.

Актуальность исследований

На разных этапах жизненного цикла прибора задачи технической диагностики имеют различную актуальность. Для этапов разработки и изготовления важными являются проверка исправности прибора, а также локализация неисправностей при их наличии. При этом требуются высокие показатели полноты контроля и глубины поиска

дефектов. Для этапа эксплуатации прибора более актуальной становится задача проверки правильности функционирования. При этом уже не столь важно, какими будут показатели полноты и глубины. Здесь на первый план выходит один из основных надежностных показателей, такой как время восстановления. Как правило, это время включает в себя время обнаружения сбоя (или отказа) и время, необходимое для его парирования (непосредственно восстановления).

Решение вопросов технической диагностики, надежности и отказоустойчивости оказывается, таким образом, тесно связано между собой и позволяет обеспечить поддержание правильности функционирования прибора в процессе эксплуатации. Актуальность исследований в этом направлении обусловлена тем, что нарушение правильности функционирования может иметь различную природу (старение элементов, воздействие неблагоприятных внешних факторов, ошибки входных данных) и приводить к повышению техногенного риска.

Теоретическая часть

Проведенные в ходе разработки исследования показали, что для обеспечения требуемой надежности вычислитель должен быть построен по трехканальной схеме с «холодным» резервом. Конструктивно вычислитель имеет в составе трехканальные модуль вторичного питания и модуль управления резервированием, а также три одноканальных модуля вычислительных. В конкретный момент времени напряжение питания подано только на один модуль вычислительный, два других остаются незапитанными (находятся в «холодном» резерве).

Вычислитель построен на основе четырех СБИС многопортового коммутатора SpaceWire 1907BM056 разработки НИИСИ РАН [1]: по одной СБИС в каждом канале резервирования используется в качестве процессора (на модуле вычислительном) и одна СБИС с отключенным процессорным ядром используется в качестве коммутатора портов SpaceWire и контроллера памяти (на модуле управления резервированием).

Вычислитель взаимодействует с другими элементами системы автоматического управления по мультиплексному каналу передачи данных (МКПД) согласно ГОСТ Р 52070 [2].

Поддержание правильности функционирования вычислителя в процессе эксплуатации заключается в циклической самодиагностике и парировании сбоев или отказов при их выявлении. Данный подход подразумевает использование информационной, временной и структурной избыточности.

Применение информационной избыточности осуществляется за счет выделения в организации оперативной памяти дополнительной области для хранения корректирующих кодов, которые заполняются и проверяются встроенным в СБИС 1907BM056 контроллером памяти. При этом самодиагностика и парирование одиночных сбоев или отказов осуществляется автоматически без участия процессора. При обнаружении двойных ошибок контроллер памяти может сигнализировать об этом с помощью прерывания.

Временная избыточность подразумевает возможность кроме выполнения штатной задачи запускать циклически программу самодиагностику прибора.

Структурная избыточность (три канала) дает шанс выполнить штатную задачу после отказа работающего канала путем переключения на резервный с восстановлением промежуточных результатов.

Для обнаружения сбоев или отказов работающий в данный момент процессор циклически проводит самодиагностику модуля вычислительного и интерфейсной части вычислителя. Самодиагностика заключается в тестовых обращениях к используемым устройствам и участкам памяти, а также в выполнении некой эталонной задачи, результат решения которой заранее известен. Результаты самодиагностики и промежуточные результаты вычислений сохраняются в специализированное ОЗУ (СпецОЗУ) с

повышенным уровнем бесбойной работы и неотключаемым напряжением питания на модуле управления резервированием. Связь между модулями организована по интерфейсу SpaceWire [3], причем обращения к СпецОЗУ реализуется через протокол удаленного доступа к памяти RMAP [4]. По результатам самодиагностики формируется слово состояния вычислителя. Биты ошибок в слове состояния устанавливаются в следующих случаях:

- при неудачном завершении теста доступа к СпецОЗУ в процессе инициализации,
- при неудачном тесте буферного ОЗУ мультиплексного канала перед формированием сообщения,
- при несовпадении окончательного результата эталонной задачи с заранее известным,
- при системных исключениях процессора (ошибки шины, обращения по невыровненным адресам, недопустимые инструкции).

При успешном завершении цикла самодиагностики процессор формирует импульс сигнала исправности. При наличии в статусном слове вычислителя установленных битов ошибок сигнал исправности сформирован не будет и программа войдет в цикл для ожидания срабатывания модуля управления резервированием.

Модуль управления резервированием работает по принципу сторожевого таймера: если в течение заданного интервала времени от активного модуля вычислительного не поступит импульс сигнала исправности, напряжение питания с него будет снято, а на один из резервных вычислительных модулей подано. Структурная схема управления резервированием приведена в [5].

Обобщенная блок-схема алгоритма программного обеспечения резервированного вычислителя показана на рисунке 2.

Когда модуль вычислительный становится активным после поступления на процессор сигнала сброса, он должен как-то определить: произошло ли начальное включение питания и необходимо начать решение штатной задачи с исходных данных или же произошло переключение на резервный канал и необходимо продолжить решение штатной задачи с неких сохраненных в СпецОЗУ промежуточных результатов. Для этого при выходе процессора из сброса анализируется некая сигнатура в СпецОЗУ, которая содержит признак (и количество) переключения на резерв. Таким образом, условия сброса анализируются, как показано на рисунке 3.

Методика исследований

Для исследования работоспособности вычислителя с рассмотренными алгоритмами было разработано программное обеспечение на языке С (с инициализацией и обработкой исключений на ассемблере архитектуры «КОМДИВ»), которое прошивалось в ПЗУ каждого модуля вычислительного. В качестве эталонной задачи был выбран поиск максимального значения двумерного массива данных (хранящегося в ПЗУ). Двумерный массив данных был смоделирован на инструментальной ЭВМ в среде MATLAB. Локальные максимумы массива по строкам сохранялись в СпецОЗУ в качестве промежуточных результатов.

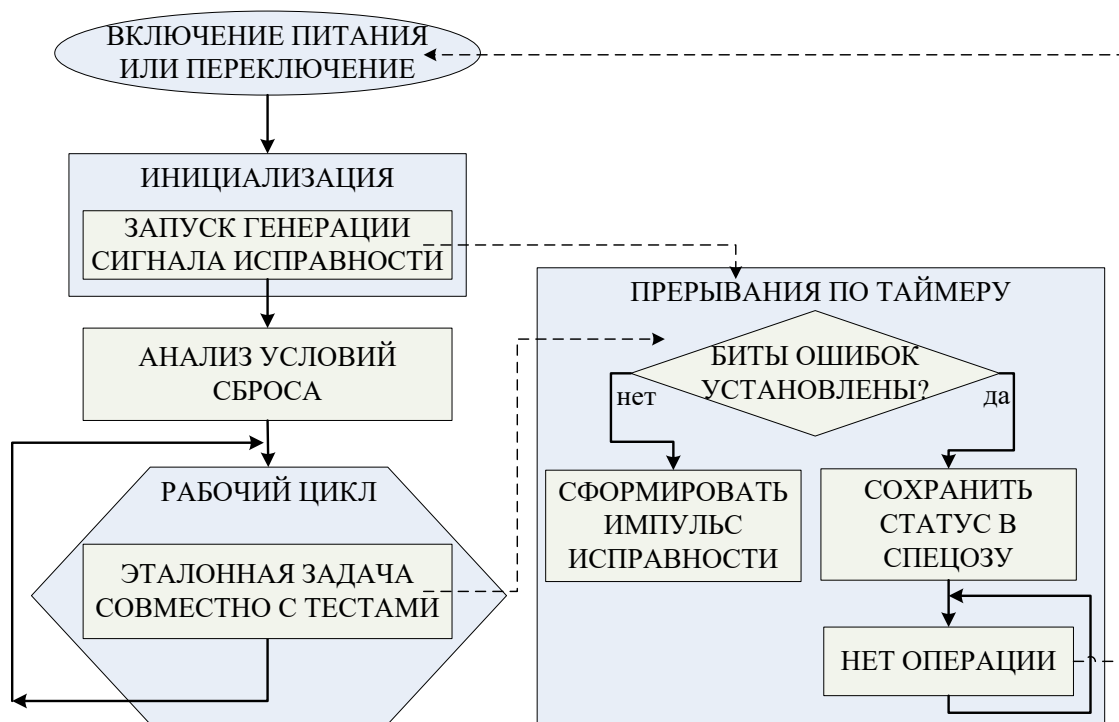


Рис. 2. Блок-схема алгоритма ПО резервированного вычислителя

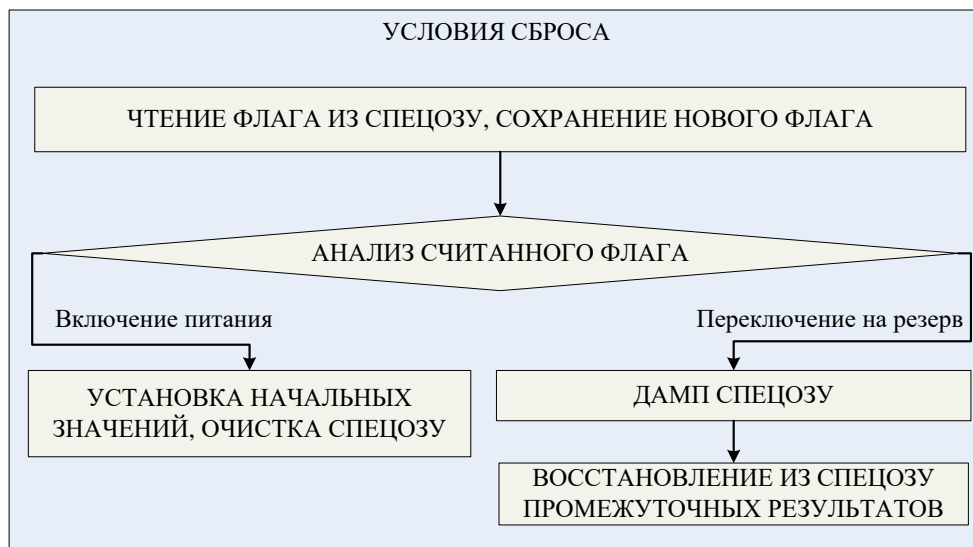


Рис. 3. Блок-схема анализа условий сброса

Для наблюдения за работоспособностью вычислителя по МКПД на инструментальной ЭВМ была разработана программа `mkpd.exe`. Программа взаимодействует с платой информационного обмена по МКПД TE2-PCI разработки ЗАО «Элкус», которая используется в качестве оконечного устройства и принимает данные от макета вычислителя, работающего в режиме контроллера шины. Программа `mkpd.exe` написана на языке C++ в среде Qt Creator и имеет графический интерфейс пользователя, который показан на рисунке 4. Особенностью данной программы является то, что в графический интерфейс выводятся принимаемые из МКПД данные как в шестнадцатеричном виде, так и интерпретированные как текст. Программа автоматически фиксирует время прихода каждого сообщения по МКПД с точностью до десятков миллисекунд. Все сообщения с метками времени сохраняются в журнал работы программы.

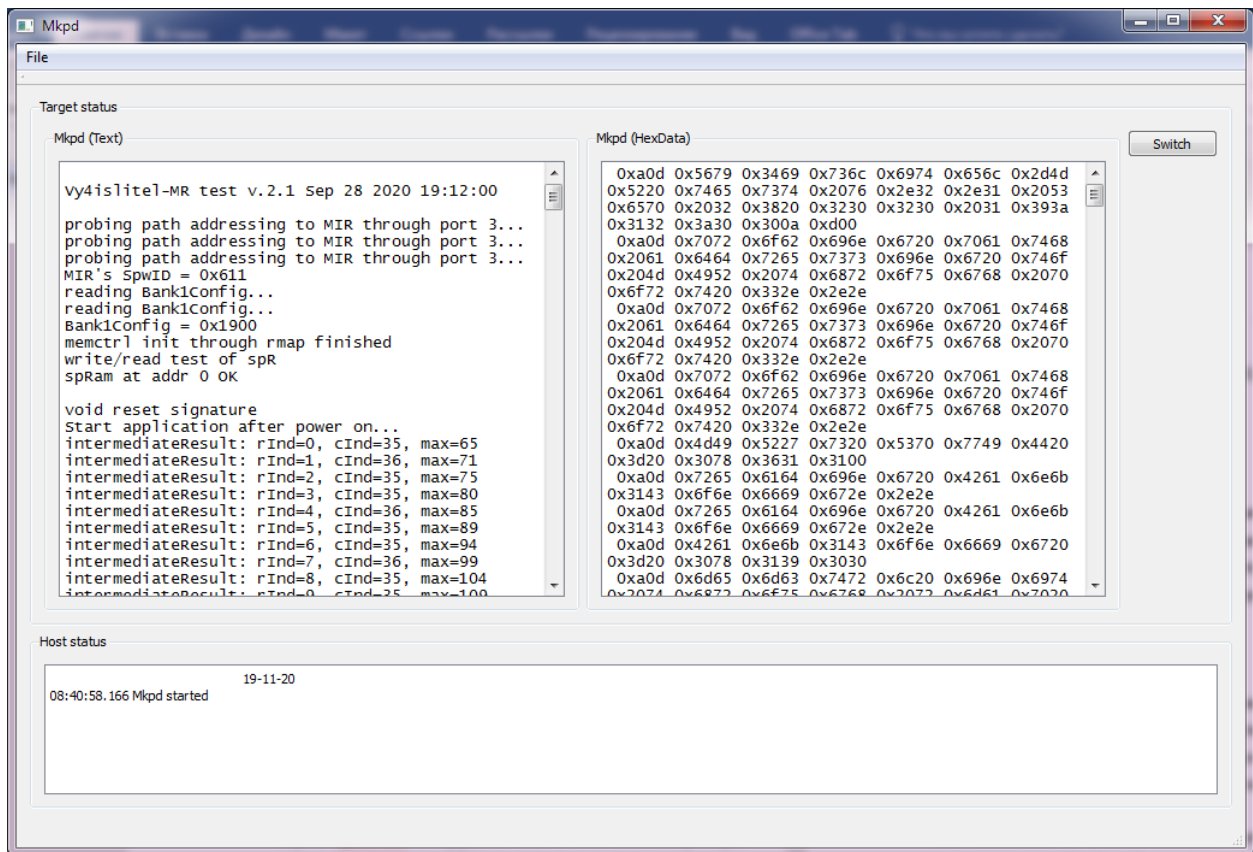


Рис. 4. Программа наблюдения за обменом по МКПД

В процессе отладки программного обеспечения вычислителя имитировались ошибки устройств с проверкой установки в слове состояния соответствующих битов.

Сбой активного вычислительного модуля имитировался подачей на него сигнала сброса. При этом схема управления резервированием фиксировала прекращение импульсов сигнала исправности, снимала с активного модуля напряжение питания и подавала напряжение на резервный.

Результаты исследований

В ходе проведения исследований отлажены входящие в состав программы самодиагностики тесты устройств.

Разработаны и отлажены функции записи и чтения данных СпецОЗУ модуля управления резервированием по интерфейсу SpaceWire через протокол удаленного доступа к памяти RMAP.

Проведена проверка работоспособности модуля управления резервированием при имитации сбоев модулей вычислительных с помощью ручного сброса.

Экспериментально подтверждено восстановление вычислений на резервных модулях с сохраненных в СпецОЗУ промежуточных результатов.

Библиографический список

1. Микросхемы интегральные 1907VM056, 1907VM05H4. Указания по применению. – М.: НИИСИ РАН, 2019. – 380 с.
2. ГОСТ Р 52070-2003. Интерфейс магистральный последовательный системы электронных модулей. Общие требования. – М.: Госстандарт России, 2013. – 26 с.
3. ECSS-E-ST-50-12C. SpaceWire – Links, nodes, routers and networks. – European Cooperation or Space Standardization (ECSS), 2008. – 129 p.
4. ECSS-E-ST-50-52C. SpaceWire – Remote memory access protocol. – European Cooperation or Space Standardization (ECSS), 2010. – 109 p.
5. Резервированная многоканальная вычислительная система: пат. 2527191 Российская Федерация / Русанов В.Н., Киселев А.Ю., Сильянов Н.В., заявитель и патентообладатель ФГУП «ФНПЦ НИИСИ им. Ю.Е. Седакова». – № 2011135121; опубл. 07.07.2014.

N.V. Sil'yanov

SELF-DIAGNOSIS AND INTERMEDIATE RESULTS RESTORING FOR COMPUTER BASED ON 1907VM056 SPACEWIRE COMMUTATOR

Branch of the RFNC-VNIIEF “NIIS n.a. Yu.Ye. Sedakov”, Nizhniy Novgorod

Three-channel cold redundant computer based on 1907VM056 SpaceWire commutator is shown. Algorithms of supporting correct operation of computer including self-diagnosis, redundancy management and restoring from specialized RAM intermediate results are developed. Interaction of computer modules through SpaceWire ports with remote memory access by RMAP-protocol is confirmed.

Keywords: self-diagnosis, restoring, SpaceWire, RMAP, 1907VM056

АВТОМАТНАЯ МОДЕЛЬ ПРОЦЕССА УСТАНОВКИ СРЕДЫ ИСПОЛНЕНИЯ MPI ПРОГРАММ НА ВЫЧИСЛИТЕЛЬНОМ КЛАСТЕРЕ

Нижегородский государственный технический университет им. Р.Е. Алексеева

Целью работы является повышение скорости установки и настройки системного программного обеспечения вычислительных Linux кластеров. Авторами предложена событийная модель управления процессами выполнения скриптов системного администрирования, обеспечивающая развертывание среды исполнения параллельных программ на основе недетерминированного конечного автомата. Приводится результат тестирования реализации событийной модели на примере платформы MPI.

Ключевые слова: конечный автомат, вычислительный кластер, MPI, операционная система, мониторинг, системное администрирование.

Введение

Бурное развитие методов и средств искусственного интеллекта, направленных на решение постоянно усложняющихся задач, требует создание эффективных алгоритмов обработки больших массивов данных. Одним из важных, приоритетных и одновременно ресурсоемких этапов построения систем AI является их обучение, требующее реализации и выполнения алгоритмического обеспечения на различных аппаратно-программных платформах, причем изменение требований к целевой системе в процессе ее развития может привести к необходимости обеспечения переносимости алгоритмов на новые компоненты вычислительных систем. В таких условиях особый интерес представляет концепция технической самоорганизации [1,2] и, в частности, принципа самоконфигурируемости [2], а также его реализация при построении технологии создания масштабируемых, расширяемых, самоконфигурируемых вычислительных кластеров.

В работе более подробно рассматривается процесс автоматического развертывания программной составляющей платформы кластера. Предлагаемые решения проиллюстрированы на примере спецификации MPI [3] передачи сообщений, широко используемой на параллельных компьютерах (кластерах), в основном на масштабируемых (SPCs) с распределенной памятью и на сети рабочих станций.

Основная цель MPI - достигнуть переносимость разрабатываемого ПО для разных параллельных компьютеров (в пределах кластера). Разработанный код также может выполняться как на группе персональных машин, входящих в кластер, так и как множество процессов на одной станции. Это, в частности, дает возможность отлаживать программы на персональном компьютере.

Параллельные и последовательные вычисления

Существует две модели программирования: последовательная и параллельная. Традиционная ЭВМ была последовательной, т.е. в определенный момент времени только над одним операндом выполнялась только одна операция.

Основная же концепция параллельных вычислений, может быть реализована двумя и более процессами обработки данных. Программа, использующая параллельные вычисления, характеризуется как набор взаимодействующих вычислительных процессов.

Параллельные программы на вычислительном кластере работают в модели передачи сообщений (message passing), то есть программа разделена на множество процессов, каждый из которых работает на своем процессоре и имеет свое адресное пространство. При этом прямой доступ к памяти другого процесса невозможен, с помощью операций приема (receive) и отправки (send) сообщений возможен обмен данными между процессами.

Такую модель легко реализовать с помощью стандартного интерфейса Message Passing Interface (MPI) реализующий обмен сообщениями между процессами, выполняющими одну задачу. MPI является наиболее распространенным стандартом интерфейса обмена данными при организации параллельных вычислений, существуют его реализации для большого числа компьютерных платформ. Этот стандарт также используется при разработке программ для кластеров и суперкомпьютеров. Как правило, MPI-программы построены по модели SPMD (одна программа - много данных), то есть для всех процессов имеется только один код программы, а различные процессы хранят различные данные и выполняют свои действия в зависимости от порядкового номера процесса.

Событийная модель автоматического развертывания и реконфигурирования MPI-кластера

Предложена событийная модель [4,5] процесса развертывания и конфигурирования платформы MPI, формализованная в виде недетерминированного конечного автомата (НКА) [4,5]. Модель описывается шестью элементами:

$$M = \langle Q, X, Y, \Delta, q_0, Q_F \rangle \quad (1)$$

где Q – множество состояний модели; X – множество входных сигналов (входной алфавит); Y – множество выходных сигналов (выходной алфавит); Δ – множество переходов: $Q \times X \rightarrow Q$; q_0 – начальное состояние системы; Q_F – множество финальных состояний.

На рис. 1 представлена схема модели НКА для выполнения поставленной задачи.

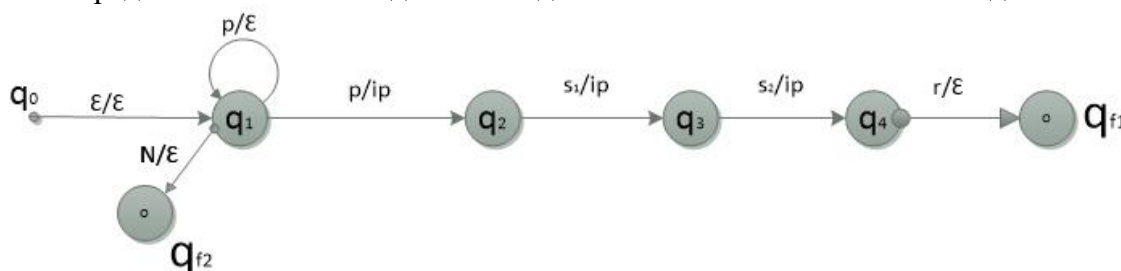


Рис. 1. Схема модели НКА.

Описание элементов событийной модели развертывания программной платформы представлено в табл. 1.

Множество входных сигналов $X = \{\varepsilon, p, s_1, s_2, r, N\}$ представлено следующими элементами: ε – переход к следующему состоянию, не зависящий от входящих и исходящих элементов; p – сигнал компьютера о готовности к работе в сети; s_1, s_2 – сигналы о завершении установки и развертывании пакетов ssh и mpi соответственно; r – результат работы тестовой программы; N – количество компьютеров, указанное заранее

Множество выходных сигналов $Y = \{\varepsilon, ip\}$ содержит пустой элемент и ip-адрес компьютера, который на начальном этапе предоставляется системой мониторинга и, далее, при передаче между состояниями автомата служит идентификатором машины, на котором запускаются скрипты.

Таблица 1

Элементы событийной модели

№ пп	Событие	Входящие	Исходящие	Комментарий
1	q_0	-	-	Состояние включения сервера и узлов кластера
2	q_1	-	Сигнал о готовности узла кластера к работе	Каждый включившийся узел посылает сигнал кластеру о готовности к работе. При достижении числа узлов равное n сервер формирует файл конфигурации
3	q_2	Адреса узлов в сети	Отправка команды на установку <code>mpi</code> с сервера	Запуск <code>ssh</code> скрипта с флагами установки
4	q_3	Файл конфигурации <code>machine.txt</code> с <code>ip</code> -адресацией	Запуск <code>mpi</code>	Установка и запуск <code>mpi</code>
5	q_4	Запуск тестовой программы, полученной от сервера	Результат работы тестовой программы	Работа тестовой программы
6	q_{f_1}	Получение результатов	-	Финальное состояние после запуска программ на узлах кластера
7	q_{f_2}	Количество работающих компьютеров в сети	-	Состояние кластера

Исследование событийной модели автоматического развертывания и реконфигурирования MPI-платформы на примере учебного вычислительного кластера

Был реализован программный прототип и пакет скриптов системного администрирования, автоматизирующих задачи установки и настройки необходимых модулей, устанавливаемых динамически на узлах локальной вычислительной сети, которые, посредством использования `ringi` сформированного файла с IP-адресами, высылают уведомления о готовности к работе.

Для обеспечения доступа к ресурсам (консоли как для Linux, так и для Windows) других компьютеров на локальном уровне использовался протокол прикладного уровня SSH, установка и настройка которого выполнялась за несколько шагов:

На сервере предварительно настроен статический IP-адрес.

Для демона `SSHD` настраиваются параметры доступа, чтобы исключить стороннее вмешательство.

Поскольку `ssh` сервер работает на 22 порту по умолчанию, то в целях повышения уровня безопасности в рамках обработчика состояния настройки `ssh` выполняются скрипты, которые:

меняют адрес порта на неиспользуемый системой;

запрещают доступ для пользователя `root`, и создают отдельного пользователя с паролем доступом.

На узлах, завершивших настройку ssh, и посылавших уведомление об этом вызываются скрипты, загружающие и устанавливающие пакеты MPI-платформы.

Тестирование MPI-кластера

Для доступа к терминалу Linux нужно набрать адрес, к которому необходимо выполнить подключение. Подключение производится к главному компьютеру кластера (серверу), с которого будет запускаться mpi программа.

В стандарте MPI описан интерфейс передачи сообщений, который может поддерживаться как платформой, так и приложением пользователя.

На сервер отправляется архив программы запуска с помощью ssh. Команда отправки:

```
Scp -P 2222 /home/student/mpi/Filestudent@192.168.15.59:/home/student/mpi
```

Вместо Fileуказывается имя пересылаемого файла.

На следующем этапе собирается программа, лежащая в папке /home/student/mpi/pi_mpi.c:

```
Mpicc -o /home/student/mpi/prog /home/student/mpi/pi_mpi.c
```

Для запуска выполнения программы на сервере создается файл конфигурации(machine), в котором будут указаны адреса компьютеров, выполняющих программу. Компьютер отправляет на сервер ring, показывая, что он готов к работе.

```
student@slave5422pc3:~$ mpirun -machinefile /home/student/mpi/machine -np 1 /home/student/mpi/prog
Process 0 of 1 is on slave5422pc2
Enter the number of intervals: (0 quits) 5
PI is approximately 3.1449258640033277, Error is 0.0033332104135346
wall clock time =0.000057
Enter the number of intervals: (0 quits) 9
PI is approximately 3.1426214565576127, Error is 0.0010288029678196
wall clock time =0.000033
Enter the number of intervals: (0 quits) 7
PI is approximately 3.1432933175274678, Error is 0.0017006639376747
wall clock time =0.000033
Enter the number of intervals: (0 quits) 1
PI is approximately 3.2000000000000002, Error is 0.0584073464102071
wall clock time =0.000033
Enter the number of intervals: (0 quits) █
```

Рис. 2. Тестовый запуск программы.

Для примера используется только сервер и один компьютер кластера, чей IP-адрес (192.168.15.2) занесен в файл конфигурации. Далее сформированный файл /home/student/mpi/prog переносится на машину 192.168.15.2.

Результаты проведенного тестирования, свидетельствующие о корректном выполнении программы на кластере, приведены на рис. 2.

Заключение

В ходе выполнения работы построена событийная модель на основе недетерминированного конечного автомата процессов развертывания среды исполнения параллельных программ. В качестве практической реализации настроен кластер для тестирования программ, а также рассмотрена методика автоматической установки, настройки и реконфигурирования MPI-кластера на базе операционной системы Linux. Для обеспечения параллельности процессов использован протокол mpi.

В дальнейших исследованиях планируется обеспечить модифицирование событийной модели в режиме реального времени для увеличения мощности кластера «на лету» при подключении новых вычислительных узлов.

Библиографический список

1. Жевнерчук Д.В. Применение методов теории самоорганизации в задачах управления профилированием и конфигурированием вычислительных систем [Текст] / Жевнерчук Д.В., Кондратьев В.В. // Доклады академии наук, том 459, №4, 2014. – С. 409-412
2. Жевнерчук Д.В. Обобщенный метод синтеза многокомпонентных интероперабельных структур на основе онтологии и недетерминированного конечного автомата // Информационные технологии. – №2 – том 25 – 2019 – С. 67-74.
3. Корнеев В.Д. Параллельное программирование в MPI. - Москва-Ижевск: Институт компьютерных исследований, 2003. - 304 с.
4. Поликарпова Н.И., Шалыто А.А. Автоматное программирование. — СПб.: Питер, 2009. — 176 с. — ISBN 978-5-388-00692-9.
5. Вашкевич Н. П. Недетерминированные автоматы и их использование для реализации систем параллельной обработки информации [Текст]: монография / Н. П. Вашкевич, Р. А. Бикташев; "Пензенский государственный университет" (ПГУ). – Пенза: Изд-во ПГУ, 2016. - 391 с.

E. S. Anisimova, E. V. Tsybulskaya, V.P. Zelenskii, D. V. Zhevnerchuk

STATE MACHINE OF THE PROCESS OF INSTALLING THE MPI PROGRAM EXECUTION ENVIRONMENT ON A COMPUTING CLUSTER

Nizhny Novgorod state technical university n.a. R.E. Alekseev

Purpose: increasing the speed of installation and configuration of system software for Linux computing clusters.

Design/methodology/approach: state machine, mpi.

Findings: the authors proposed an event-based model for managing the processes of executing system administration scripts, which provides the deployment of a parallel program execution environment based on a nondeterministic state machine.

Research limitations/implications: criteria can be used for systems of various configuration and difficulty.

Originality/value: the model can be used to speed up the process of computing big data on a computing cluster of various capacities.

Keywords: state machine, computing cluster, MPI, operating system, monitoring, system administration.

МЕТОДИКА ПОСТРОЕНИЯ ПАРАЛЛЕЛЬНОГО ПРОГРАММНОГО КОДА НА ОСНОВЕ АВТОМАТНЫХ МОДЕЛЕЙ

Нижегородский государственный технический университет им. Р.Е. Алексеева

Целью данной работы является разработка программного обеспечения для создания автоматных моделей управления взаимодействующими процессами и алгоритмов генерации программного кода на их основе. Полученные результаты проиллюстрированы на примере классической задачи межпроцессного взаимодействия “Читатели и писатели”.

Ключевые слова: графы, МРІ, недетерминированные конечные автоматы, алгоритмы управления взаимодействующими процессами

Введение

Повышение производительности вычислительных машин и систем автоматики на их основе во многом определяется эффективностью организации логического управления в системах, которые учитывают реализацию параллелизма [1]. При создании таких систем одной из актуальных задач является разработка формальных методов описания алгоритмов управления параллельными взаимодействующими процессами, их эквивалентных преобразований и методов их структурной реализации на разных уровнях.

Такой подход к формализации сложных параллельных алгоритмов управления с взаимодействующими ветвями позволяет создать набор методов и средств для компактного формального описания алгоритмов с последующим построением возможных реализаций с учетом требований повышения производительности, надежности и снижения требований к аппаратным ресурсам.

Событийные модели, основанные на использовании концепции параллелизма и недетерминизма, являются простыми и универсальными. Они могут быть использованы при формализации многопроцессных систем широкого класса.

К сожалению, известное инструментальное программное обеспечение, реализующее автоматный подход к генерации кода параллельных программных систем, как правило, ориентировано на конкретные платформы и не может быть использовано по схеме: “1 модель – n программ”. С другой стороны, перевод граф-схем автоматов (ГСА) вручную занимает достаточно количество времени и требует привлечения специалистов, обладающих компетенциями в этой области. Поэтому целью данной работы является создание программного комплекса построения автоматных моделей и алгоритмов управления взаимодействующими процессами по граф-схеме автомата (автоматов), построение спецификации описания граф-моделей на основе кроссплатформенных форматов и поведение тестирования построенных с помощью этого инструмента моделей на примере автоматического построения кода МРІ программ [2].

Проблематика и задачи

Известные отечественные научные школы, развивающие автоматный подход в области информационно-телекоммуникационных технологий и вычислительной техники представлены следующими направлениями:

Организация логического управления в системах с распараллеливанием на всех уровнях управления [3]. Разработка систем параллельной обработки информации на основе недетерминированных конечных автоматов.

Автоматное программирование [4]. Switch-технология – технология разработки систем логического управления на базе конечных автоматов, охватывающая процесс спецификации, проектирования, реализации, отладки, верификации.

Дискретно-событийное и мультиагентное моделирование [5].

Методика генерации MPI кода на основе НДКА

На рис.1 изображена схема основных компонентов системы. Для начала работы с программой пользователь должен создать пустую ГСА (блок 2) или загрузить уже существующую ГСА задачи (блок 1). Система создаст или загрузит ГСА и пользователь сможет приступить к работе над ГСА (блок 3). На схеме «ГСА задачи» показывает что у пользователя должна быть ГСА задачи в удобном ему виде.

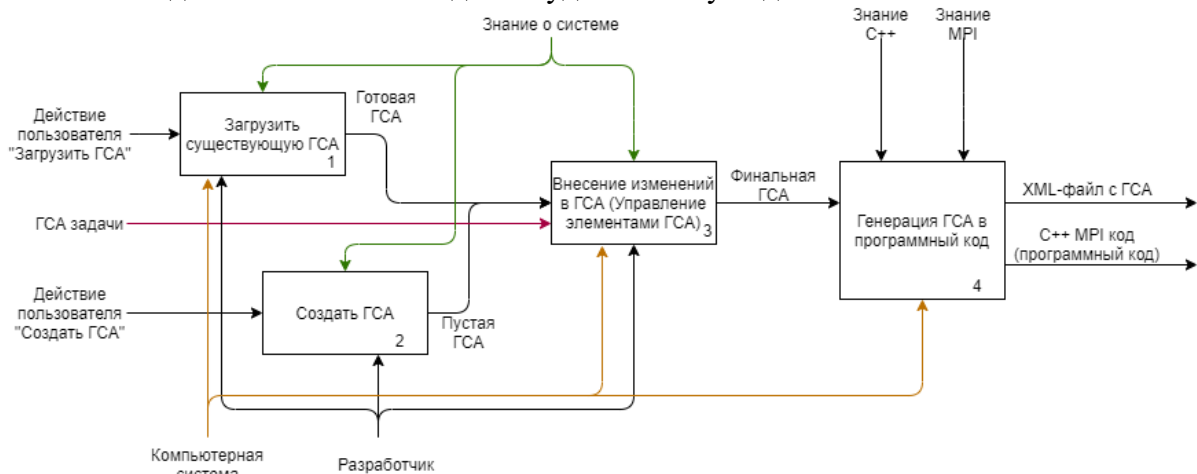


Рис. 1. Схема генерации программного кода из ГСА задачи

На данном этапе (блок 3) пользователь может изменять ГСА, вводить в программный код в вершины ГСА. Как только пользователь пожелает ГСА перевести в программный код, то пользователь вызывает соответствующее действие и финальная ГСА отправляется в генератор программного кода (блок 4). Генератор программного кода анализируя ГСА строит MPI-код в виде скелета. Когда скелет кода MPI будет построен, встраивается пользовательский код каждой вершины в скелет кода MPI. Когда код будет скомпилирован, программа сохранит XML-файл ГСА задачи и C++ MPI код на жесткий диск.

Экспериментальная часть

Рассмотрим применение теоретических результатов на примере задачи читателей-писателей. На рис. 2 представлена автоматная модель, отражающая особенности межпроцессного взаимодействия.

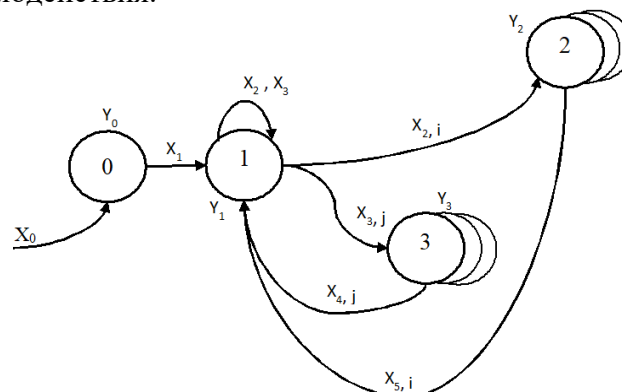


Рис. 2. Система управления межпроцессным взаимодействием “Читатели и писатели”

В табл. 1 представлено описание элементов построенной автоматной модели.

Таблица 1

Пояснение к ГСА задачи (рис. 4)

Событие	Назначение события
0	Подготовка к работе (Настройка автомата)
1	Рабочее состояние (Ведущая логика собрана именно здесь)
2, 3	Группы событий (ведомые потоки или читатели/писатели)

Управляющее слово	Назначение управляющего слова
i, j	Выбор конкретного потока/клиента
X_0	Старт автомата
X_1	Переход в активное состояние автомата (старт основной работы автомата)
X_2 / X_3	Читателю/Писателю разрешен вход в «книжный шкаф»
X_4 / X_5	Писатель/Читатель освобождает «книжный шкаф»
Y_0	подготовка к работе
Y_1	определение приоритета на коллекцию, контроль критической секции
Y_2 / Y_3	Читатель/Писатель берет/приносит книгу

Введем систему идентификации потоков. Поскольку существует всего два типа клиентов – читатели и писатели, то индекс каждого потока может быть использован для определения типа: читатель или писатель. Для этого вычислим остаток от деления значения индекса на 2. Если получили нечетное значение, то будем считать этот поток писателем, если четное – писателем.

Программный макет

Программный макет был собран по схеме рис. 1. На главном окне макета (рис. 3) есть все необходимые пользователю функции: создавать, загружать, изменять, удалять элементы ГСА и саму ГСА, а также её сохранять и компилировать в программный код. Пользователь может выбрать вершину ГСА в программе и в окне вершины (рис. 4) задать программный код, который будет обрабатываться выбранной вершиной.

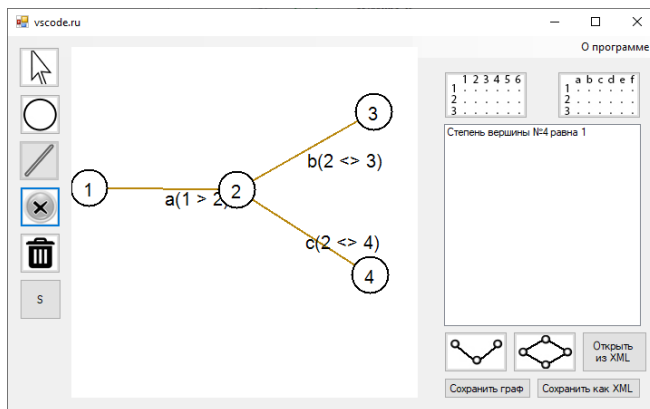


Рис. 3. Главный вид программы

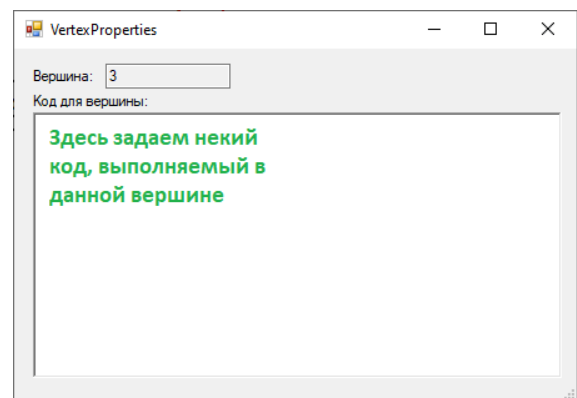


Рис.4. Окно вершины

Также программная реализация обладает функцией импорта/экспорта построенных автоматных моделей в/из формата XML.

Структура XML-файла

Рассмотрим структуру XML файла, сохраняемый данной программой. Описание тэгов представлено в таблице 2.

Таблица 2

Описание xml-тэгов для хранения ГСА

Тэг	Описание тэга
<Graph>	Root-элемент, в нем хранятся все элементы ГСА
<V>	Набор вершин. Состоит из коллекции Vertex тэгов.
<E>	Набор ребер. Состоит из коллекции Edge тэгов.
<Vertex>	Вершина ГСА. Состоит из x, y, name и pcode тэгов
<Edge>	Ребро ГСА. Связывает 2 вершины между собой. Состоит из v1, v2, rv1, rv2, name и generateName тэгов.
<x>, <y>	Теги координат вершины ГСА. Данные вершины нужны для отрисовки ГСА в программном макете.
<name>	в Vertex – название/индекс вершины в Edge – название ребра с указанием направлений
<pcode>	Программный код для обработчика вершины
<generatedName>	Название ребра без указания направлений

На рис. 5 представлен xml-файл с описанием фрагмента ГСА, формализующей межпроцессное взаимодействие процессов-читателей и процессов-писателей.

```

<?xml version="1.0" ?>
<Graph>
  <V>
    <Vertex>
      <x>16</x>
      <y>168</y>
      <name>1</name>
      <pcode />
    </Vertex>
    ...
  </V>
  <E>
    <Edge>
      <v1>0</v1>
      <v2>1</v2>
      <rv1>-1</rv1>
      <rv2>-1</rv2>
      <name>a(1 &gt; 2)</name>
      <generatedName>a</generatedName>
    </Edge>
    ...
  </E>
</Graph>

```

Рис. 5. Структура XML-файла

Маркерами выделены блоки xml, с которыми связаны пояснения:

- координаты вершины на холсте (маркер 1)
- метка вершины (маркер 2)
- программный код для вершины (маркер 3)
- блок вершины (маркер 4)
- набор (коллекция) вершин (маркер 5)
- индексы соединения вершин в прямом направлении (индексы берутся из набора V) (маркер 6)
- индексы соединения вершин в обратном направлении (если есть второе направление) (маркер 7)
- метка ребра с полным определением (маркер 8)
- метка ребра (маркер 9)

- блок ребра (маркер 10)
- набор (коллекция) ребер (маркер 11)

На рис. 6 представлен фрагмент программного кода, сгенерированного на основе построенной ГСА. Для улучшения восприятия в код были внесены комментарии.

```

1 while(1)
{
  if (RANK == 0) //системный поток
  2 {
    if (code == 0b00000000) //PREPARING, подготовка
    {
      3 //инициализация книжного шкафа, создание клиентов шкафа, распределение ролей (читатель/писатель)
      //переходим на след. вершину автомата
      code = 0b00000001;
    }
    else if (code == 0b00000001) //рабочая вершина автомата
    {
      //выборка приоритета: если у нас пустая кол-я - только писатель, если целая - только читать, остальные - как рандом скажет
      //случай 1: если шкаф пустой, выбираем только писателей и среди писателей любого
      4 //случай 2: если шкаф полон, выбираем только читателей и среди читателей любого
      //случай 3: в остальных случаях - выбираем любого
      MPI_Request request;
      MPI_Isend(&book_cabinet, 1, MPI_INT, selectedRank, 0b00000001, MPI_COMM_WORLD, &request);
      code = 0b11111110; //уходим в ожидание
    }
    else if (code == 0b11111110) //ожидающая вершина
    {
      int message; //RANK
      MPI_Status st;
      MPI_Recv(&message, 1, MPI_INT, MPI_ANY_SOURCE, MPI_ANY_TAG, MPI_COMM_WORLD, &st);
      if(st.MPI_TAG == 0b00000010) //Освобождаем крит. секцию
      {
        5 if (selectedRank == st.MPI_SOURCE)
        {
          book_cabinet = message;
          selectedRank = -1; //оброс критической секции
          code = 0b00000001; //следующий клиент
        }
        else if(st.MPI_TAG == 0b00000011) //Клиент готов
        {
          //отметить клиента свободным в коллекции клиентов
        }
      }
      else if (code == 0xFF) //выход из программы
      {
        6 //бродкастовая рассылка некого сигнала для завершения работы
      }
      7 Sleep(20);
    }
    else
    {
      int message;
      MPI_Status st;
      MPI_Recv(&message, 1, MPI_INT, 0, MPI_ANY_TAG, MPI_COMM_WORLD, &st);
      if (st.MPI_TAG == 0xFF)
      {
        8 //завершение работы
      }
      else if (st.MPI_TAG == 0b00000001) {
        9 book_cabinet = message;
        if (channel == 0) //читатель/читатели
        {
          //клиент берет книгу, снимает блокировку, освобождается
        }
        else if (channel == 1) //писатель/писатели
        {
          //писатель кладет книгу, снимает блокировку, освобождается
        }
      }
    }
  }
}
MPI_Finalize();
return 0;
}

```

Рис. 6. Листинг сгенерированного кода

Маркерами выделены блоки сгенерированного кода, с которыми связаны пояснения:

- бесконечный цикл является основой данной технологией. Он позволяет запускать вершину снова и снова, тем самым мы не теряем связь с вершиной на всем протяжении работы программы (маркер 1),
- блок системного потока. Здесь вся логика по управлению переходами состояний автомата (маркер 2),
- код обработчика вершины 0 (маркер 3),
- код обработчика вершины 1 (маркер 4),
- код обработчика вершины 1 в режиме «Блокировка» (маркер 5),
- выход из программы (маркер 6),
- сделано ожидание в 20 мс для отладочных целей (маркер 7),
- блок обработки вершин (маркер 8),
- код обработчиков вершины 2 и 3 (маркер 9).

Заключение

Предложена методика, предназначенная для автоматического построения MPI программ на основе граф-схемы автомата задачи. Разработан прототип инструментального программного обеспечения, реализующего редактирование ГСА импортирование/экспортирование ГСА в XML формате, генерацию программного кода на C++ с поддержкой спецификации MPI. В качестве демонстрации, была выбрана задача межпроцессного взаимодействия “Читателей и писатели”.

Применение методики, в основе которой лежит теория автоматов, может помочь существенно облегчить работу программистов при разработке параллельных алгоритмы и их реализации. На текущем этапе существуют ограничения, связанные с тем, что для ряда частных случаев требуется коррекция программного кода “вручную”.

В дальнейшем планируется развитие семантического обеспечения ГСА и создание спецификации API автоматного и акторного описания систем управления взаимодействующими процессами, а также алгоритмов генерации программного кода на других платформах, поддерживающих распределенные вычисления.

Библиографический список

1. Воеводин В.В. Параллельные вычисления / В.В. Воеводин, Вл. В. Воеводин // БХВ-Петербург. – 2002. – 608 с.
2. Корнеев В.Д. Параллельное программирование в MPI / В.В. Воеводин, Вл. В. Воеводин // Москва-Ижевск: Институт компьютерных исследований. – 2003. – 304 с.
3. Вашкевич Н.П. Недетерминированные автоматы и их использование для реализации систем параллельной обработки информации [Текст]: монография / Н. П. Вашкевич, Р.А. Бикташев; Министерство образования и науки РФ, Федеральное государственное бюджетное образовательное учреждение высшего образования "Пензенский государственный университет". - Пенза: Изд-во ПГУ, 2016. - 391 с.
4. Поликарпова Н.И. Автоматное программирование / Н.И. Поликарпова, А.А. Шалыто // СПб: Питер. – 2009. – 176 с. – ISBN 978-5-388-00692-9
5. Боев В.Д. Исследование адекватности GPSS World и AnyLogic при моделировании дискретно-событийных процессов: Монография. – СПб.: ВАС. – 2011. – 404 с.

A.A. Karlyshev, D.V. Zhevnerchuk

TECHNIQUE FOR CONSTRUCTING PARALLEL PROGRAM CODE BASED ON AUTOMATONS MODELS

Nizhny Novgorod state technical university n.a. R.E. Alekseev

Purpose: software development for creating automatons models of control of interacting processes and algorithms for generating program code based on them.

Design/methodology/approach: graph theory

Findings: the authors have proposed a technique for constructing parallel program code based on automatons models.

Research limitations/implications: technique can be used for graph-schemes of automaton of task of various difficulty

Originality/value: technique can be used for creating C++ MPI program code from graph-schemes automaton of task and fragments of C++ user code for every vertex of graph-schemes automaton.

Keywords: graphs, MPI, non-deterministic finite-state automaton, algorithms of control of interacting processes

КРИТЕРИИ ЭФФЕКТИВНОСТИ РЕКОНФИГУРИРОВАНИЯ МНОГОКОМПОНЕНТНЫХ СИСТЕМ НА ОСНОВЕ ИХ СТРУКТУРНО- ТОПОЛОГИЧЕСКИХ ХАРАКТЕРИСТИК

Нижегородский государственный технический университет им. Р.Е. Алексеева

Целью работы является повышение эффективности сборки и реконfigurирования многокомпонентных систем. Авторами предложены критерии реконfigurирования систем, применение которых позволяет повысить качество структуры информационной системы за счет удаления избыточных компонентов без потери ее функциональности.

Ключевые слова: многокомпонентная система, интерфейс, критерий эффективности, структурно-топологические характеристики системы.

Введение

При проведении структурного анализа систем необходимо иметь возможность определения структурно-топологических характеристик систем и получения их количественной оценки. Целесообразность определения этих характеристик заключается в том, что на стадии проектирования системы должно быть оценены качество структуры системы и ее элементов с позиций системного подхода. В качестве основных структурно-топологических характеристик системы можно выделить следующие [1]:

1. Связность структуры – количественная характеристика, позволяющая выявить наличие обрывов в структуре и определить компоненты, не связанные с другими компонентами системы. Если, применив компонентный подход [2], представить систему в виде набора компонентов, взаимодействующих посредством стандартизированных интерфейсов [3], и затем перейти к представлению в виде графа (вершины графа – компоненты системы, дуги – интерфейсы), то связность элементов будет определяться матрицей связности $C = \|C_{ij}\|$, элементы которой определяются на основе суммарной матрицы смежности $A^S = \sum_{k=1}^n A^k$, где n – число вершин графа, $A^k = A^{k-1} \cdot A$, где A – матрица смежности графа. Элемент матрицы $A^k = \|a_{ij}^k\|$ определяет число путей длиной k от вершины i к вершине j . Таким образом элемент матрицы связности будет иметь вид:

$$c_{ij} = \begin{cases} 1, & \text{если } a_{ij}^s \geq 1; \\ 0, & \text{если } a_{ij}^s = 0. \end{cases}$$

2. Структурная избыточность – характеристика, отражающая превышение общего числа связей над минимально необходимым. Определяется следующим образом:

$R = \frac{1}{2} \cdot \frac{1}{n-1} \cdot \sum_{i=1}^n \sum_{j=1}^n a_{ij} - 1$. Для систем, обладающих избыточностью, $R > 0$, а для систем без избыточности $R = 0$.

3. Структурная компактность – количественная характеристика, отражающая близость элементов системы между собой. Определяется следующим образом:

$Q = \sum_{i=1}^n \sum_{j=1}^n d_{ij}, i \neq j$, где d_{ij} – минимальная длина пути между элементами i и j .

Критерии эффективности реконfigurирования многокомпонентных систем

В общем случае для систем сложной структуры количество компонентов может быть настолько большим, что алгоритмы, позволяющие осуществлять сборку и реконfigurирование таких многокомпонентных структур будут обладать высокой вычислительной сложностью и, как следствие, значительным временем выполнения. В связи с этим при реконfigurировании многокомпонентных систем (МКС) актуальным является снижение количества компонентов системы без потери ее функциональности, а также обнаружение и устранение ошибок, допущенных при их сборке – компонентов, для которых отсутствуют связи с другими компонентами системы, либо компонентов с нарушенным порядком следования. Сформулируем свойства компонентов МКС, а также основные показатели, связанные с МКС (табл. 1).

Таблица 1

Показатели, связанные с МКС

№ пп	Наименование	Обозначение	Описание
1	Информационный поток	ψ_m	Информационный поток на выходе интерфейса m , $\psi_m \in \Psi$, где Ψ – множество информационных потоков на выходе интерфейсов МКС
2	Компонент	c_i	$c_i \in C$, где $C = \{c_1, \dots, c_r\}$ – множество компонентов МКС
3	Переход между компонентами c_i и c_j по информационному потоку ψ	$\langle c_i, \psi \rangle \rightarrow c_j$	Существует переход по информационному потоку ψ между компонентами c_i и c_j
4	Количество переходов по информационному потоку ψ между компонентами	l	Характеризует избыточность МКС по компонентам $l = \left \bigcup_{i=1}^r \langle c_i, \psi \rangle \rightarrow c_j \right $
5	Цепочка компонентов τ	$\tau = \langle c_i, \dots, c_k \rangle$	Цепочка последовательно связанных между собой компонентов посредством входных/выходных информационных потоков $\tau \in T$, где $T = \{\tau_1, \dots, \tau_w\}$ – множество всех цепочек компонентов в рамках данной МКС, а w – их общее количество

Свойства компонентов МКС:

1. Эквивалентность компонентов ($c_1 \equiv c_2$). Эквивалентными по входным потокам данных будут считаться компоненты, обладающие одинаковыми информационными потоками на входах их входных интерфейсов ($X_{c_1} \equiv X_{c_2}$). Эквивалентными по выходным потокам данных будут считаться компоненты, обладающие одинаковыми информационными потоками на выходах их выходных интерфейсов ($Y_{c_1} \equiv Y_{c_2}$).

2. Бесплезность компонентов. Недостижимыми будут считаться компоненты, у которых отсутствуют входные интерфейсы, неразрешимыми – у которых отсутствуют выходные интерфейсы. Такие компоненты нарушают связность структуры системы и являются бесплезными, поскольку никак не сказываются на функционировании системы, не обмениваясь данными с остальными ее компонентами.

3. Упорядоченность компонентов. Если информационный поток на выходе выходного интерфейса компонента должен быть подан на вход входного интерфейса другого компонента, можно считать, что эти компоненты образуют упорядоченную цепочку компонентов.

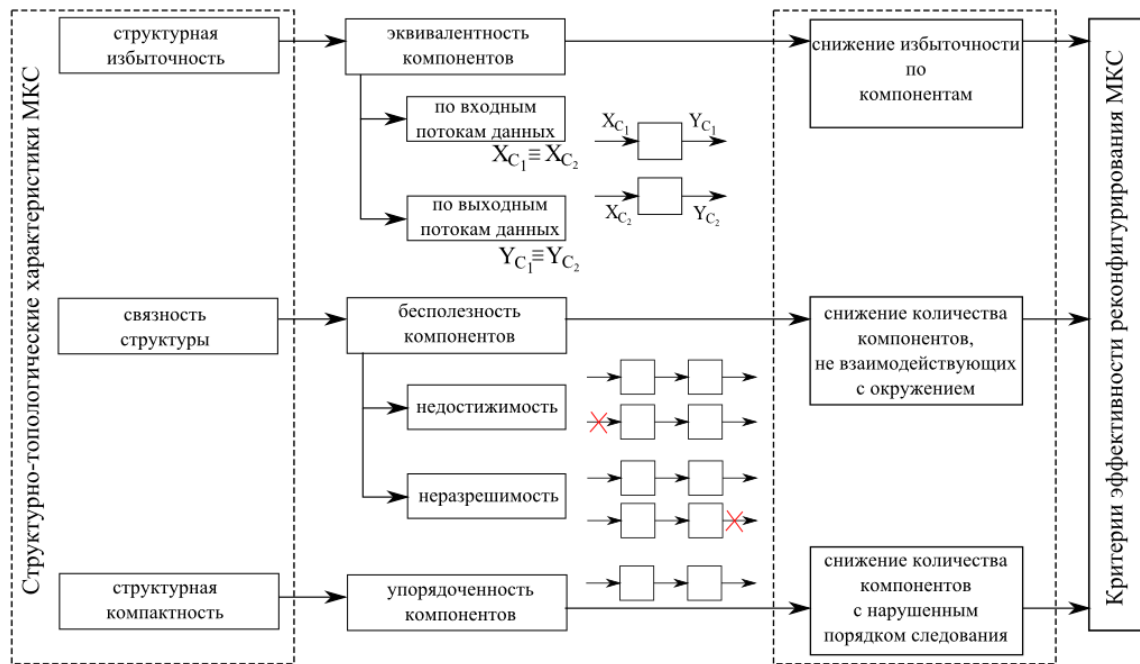


Рис. 1. Процесс формирования критериев эффективности реконфигурирования MKC на основе их структурно-топологических характеристик

Таким образом, используя введенные показатели, сформулируем критерии эффективности реконфигурирования MKC.

1. Снижение степени избыточности MKC по компонентам.

Используя введенные характеристики MKC (табл. 3.1), сформулируем критерий избыточности многокомпонентной структуры: MKC содержит избыточное количество компонентов, если существуют два и более компонента, обладающие выходными интерфейсами, выдающими такие информационные потоки ψ_m и ψ_n , что $m = n$:

$$\begin{cases} l_{\rightarrow\psi} = \min \left(\bigcup_{i=1}^r \langle c_i, \bar{\psi} \rangle \rightarrow \bar{c} \right) \\ \bar{c} \subseteq C \end{cases}$$

Данный критерий позволяет минимизировать избыточность MKC путем снижения количества компонентов, между которыми происходят переходы по одинаковым информационным потокам.

Для данного критерия показатель эффективности будет выглядеть следующим

образом: $k_1 = \left(\frac{\sum_{i=1}^q |c'_i|}{|C|} \right) \cdot 100\%$, где $\bar{c}_i \subseteq C$ – подмножество обнаруженных компонентов с

одинаковыми информационными потоками на выходе, $c'_i \subseteq \bar{c}_i$ – подмножество удаленных компонентов, а $q = |\Psi|$ – общее количество информационных потоков на выходах интерфейсов MKC.

2. Снижение количества компонентов MKC, не взаимодействующих с внешним окружением

$$\begin{aligned} l_{\rightarrow\emptyset} &= \min \left(\bigcup_{i=1}^r \langle c_i, \bar{\psi} \rangle \rightarrow \emptyset \right) \\ l_{\emptyset\rightarrow} &= \min \left(\bigcup_{i=1}^r \langle c_i, \emptyset \rangle \rightarrow c_j \right) \end{aligned}$$

Данный критерий позволяет минимизировать число компонентов МКС, не взаимодействующих с внешним окружением, то есть не связанных информационными потоками с другими компонентами (по своему входу либо выходу).

Для данного критерия показатель эффективности будет выглядеть следующим образом: $k_2 = \frac{|T_{\text{корр}}|}{|T|} \cdot 100\%$, где $T_{\text{корр}}$ – множество корректно построенных цепочек компонентов.

3. Снижение количества компонентов МКС, для которых нарушен порядок в цепочке компонентов

$$l_{\rightarrow} = \min(|\langle c_i, \psi \rangle \rightarrow c_j|), \forall i > j$$

Данный критерий позволяет повысить качество структуры многокомпонентной системы за счет обнаружения компонентов, для которых нарушен порядок в цепочке компонентов, и перестройки соответствующих им цепочек компонентов.

Для данного критерия показатель эффективности будет выглядеть следующим образом: $k_3 = \frac{|T_{\text{пер}}|}{|T|} \cdot 100\%$, где $T_{\text{пер}}$ – множество перестроенных цепочек компонентов.

Данные критерии сформулированы таким образом, что их применение не приведет к снижению функциональности системы. В рамках первого критерия определяются эквивалентные компоненты (они обладают либо на своих входах, либо на своих выходах одними и теми же информационными потоками), то есть избыточные либо дублирующиеся, поэтому их исключение из системы не скажется на ее функциональности. В рамках второго критерия выявляются компоненты, не взаимодействующие с внешним окружением, то есть у них отсутствуют либо входные, либо выходные интерфейсы, удаление этих компонентов так же не скажется на функциональности системы. По третьему критерию определяются компоненты, для которых нарушен порядок следования, то есть неправильно выстроен порядок с точки зрения взаимодействия по их интерфейсам.

Реконfigurирование имитационных моделей СМО

Полученные результаты были применены при решении задачи структурной оптимизации имитационной модели, распределенной вычислительной blockchain системы. Предложена многокомпонентная система (рис. 2), определяющая фильтр, позволяющий выявить сегмент по имитационной модели, содержащей избыточность устройств обслуживания (вычислительных узлов). В состав фильтра входят:

- генератор транзактов с выходным интерфейсом, передающим объект транзакт.
- устройство обслуживания, входной интерфейс которого принимает объект транзакт, а выходной выдает тот же объект транзакт и параметры задержки (обслуживания), включающий тип и параметры закона распределения интервала времени обслуживания.

В результате применения фильтра определялись избыточные устройства обслуживания, которые заменялись на:

- компонент инициализации, передающий в переменные объекта транзакта имя очереди, устройства, а также параметры генератора интервала времени обслуживания на устройстве.
- компонент конфигурируемого устройства обслуживания, настройка которого базируется на механизме косвенной адресации GPSS World.

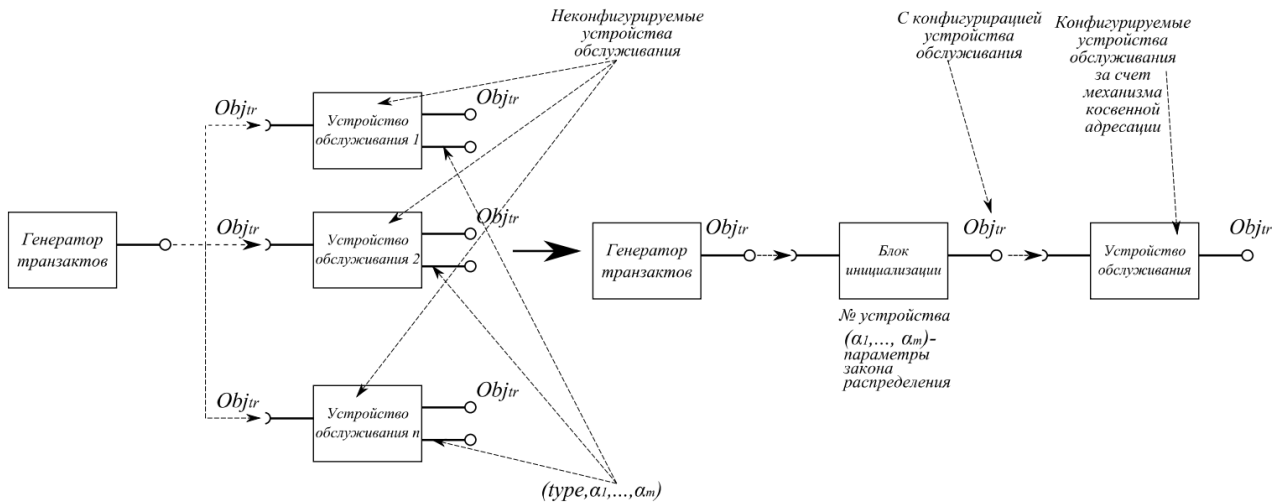


Рис. 2. Многокомпонентная система, определяющая фильтр

На рис. 3 приведены фрагменты модели расчета хэш-значений по данным, до и после реконфигурирования. Для лучшего восприятия модель содержит 4 узла, в общем случае их количество определялось 10^4

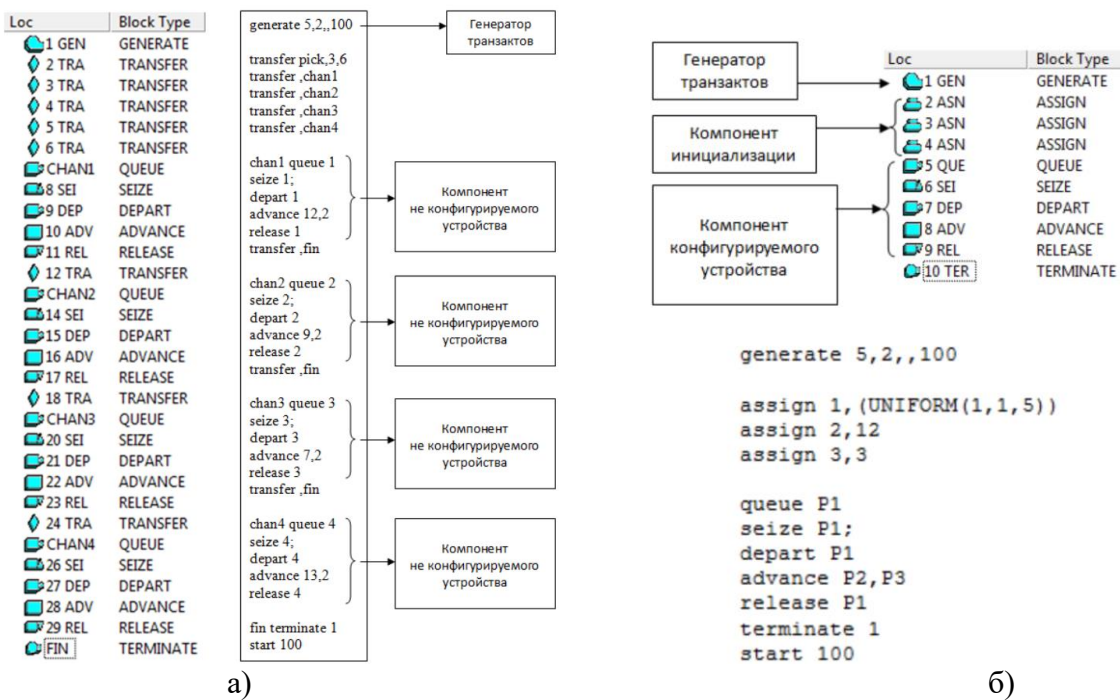


Рис. 3. Фрагмент модели расчета хэш-значений по данным а) до реконфигурирования, б) после реконфигурирования

В результате реконфигурирования моделей, размер кода был сокращен:

- до 30% от исходного, в случае, когда процесс обслуживания устройств описывается одинаковым законом;
- до 79-91% от исходного, в случае, когда процесс обслуживания устройств описывается разными законами распределения интервала обслуживания (на множестве устройств обслуживания существуют подмножества с одинаковыми законами обслуживания).

Заключение

В работе выделены показатели, связанные с многокомпонентными системами, и свойства компонентов. Сформулированы критерии эффективности реконfigurирования многокомпонентных систем и связанные с ними показатели эффективности. Приведены результаты применения полученных критериев для решения задачи структурной оптимизации модели системы массового обслуживания на примере фрагмента модели, распределенной вычислительной blockchain системы.

Библиографический список

1. Воропай Н.И. Теория систем для электроэнергетиков: учебное пособие / Н.И. Воропай // Новосибирск: Наука, Сибирская издательская фирма РАН, 2000. – 273 с.
2. Жевнерчук Д.В. Обобщенный метод синтеза многокомпонентных интероперабельных структур на основе онтологии и недетерминированного конечного автомата / Д.В. Жевнерчук // Информационные технологии. – 2019. – №2 – том 25 – С. 67-74.
3. Кулясов П.С. Моделирование и оптимизация адаптивных многокомпонентных систем на базе алгебраических структур / Д.В. Жевнерчук, П.С. Кулясов // Моделирование, оптимизация и информационные технологии. – 2020. – №1(8). Доступно по: https://moit.vivt.ru/wpcontent/uploads/2020/02/ZhevnerchukKulyasov_1_20_1.pdf DOI: 10.26102/2310-6018/2020.28.1.005.

P.S. Kulyasov, D.V. Zhevnerchuk

EFFECTIVENESS CRITERIA OF RECONFIGURING MULTICOMPONENT SYSTEMS BASED ON THEIR STRUCTURAL-TOPOLOGICAL CHARACTERISTICS

Nizhny Novgorod state technical university n.a. R.E. Alekseev

Purpose: increasing the efficiency of assembly and reconfiguration of multicomponent systems.

Design/methodology/approach: system analysis, set theory, graph theory.

Findings: the author has proposed an effectiveness criteria of reconfiguring multicomponent systems.

Research limitations/implications: criteria can be used for multicomponent systems of various configuration and difficulty.

Originality/value: criteria can be used for solving the problem of structural optimization of a queuing system model, using the example of a fragment of a model of a distributed computing blockchain system.

Keywords: multi-component system, interface, effectiveness criteria, structural-topological characteristics of system.

РЕФАКТОРИНГ МОДЕЛЕЙ ПРОГРАММНОГО ОБЕСПЕЧЕНИЯ КАК ОДИН ИЗ ЭТАПОВ ОПТИМИЗАЦИИ

Нижегородский государственный технический университет им. Р.Е. Алексеева

В работе представлен алгоритм применения алгебраических структур в рефакторинге, показана его математическая основа, а также приводится возможность его комбинации с методами рефакторинга моделей программ, основанными на паттернах проектирования и рассматривается влияние процесса рефакторинга на оптимизацию программ.

Ключевые слова: алгебраические структуры, рефакторинг, шаблоны проектирования, UML-диаграммы, оптимизация.

Введение

Рынок программных систем, применяющихся в различных областях науки и производства, в настоящее время развивается стремительными темпами, количество программ увеличивается практически ежедневно. Некоторые продукты появились относительно недавно, а некоторые продолжают существовать, разрабатываться и обрывать новым функционалом по несколько десятков лет. Поэтому одной из важных задач в этой области стала задача улучшения качества кода, его внутренней структуры, ведь он, возможно, писался десятками различных программистов в разное время и с разным опытом и требованиями к функционалу.

Ошибки, не выявленные в ходе тестирования ПО, как правило, приводят к снижению надежности и затягиванию сроков его внедрения. Поэтому актуальность разработки совершенного кода очень высока, так как она позволяет повысить его надежность. Очевидно, что такой код должен быть максимально оптимальным.

Примитивный, но правильный код, написанный программистом, во многих случаях может быть усовершенствован. Чаще всего причиной является то, что выбранный алгоритм, является шаблонным и не учитывает условия поставленной задачи, то есть транслирует языковые выражения вне зависимости от их смысла в определенные последовательности команд. Формальный алгоритм не различает особые случаи и не использует их выгод. Выбор такого подхода приводит к результатам, которые лишь отчасти отвечают требованиям экономии памяти и скорости выполнения.

Рефакторинг моделей программ

Рефакторинг программного кода — это процесс изменения его структуры таким образом, чтобы не затрагивалось внешнее поведение исходного кода. Одна из основных целей этого процесса - облегчение понимания программного кода человеком и, пусть и не всегда, оптимизация производительности программы, вследствие чего код начинает работать быстрее.

В основе рефакторинга лежит последовательность небольших преобразований программного кода, сохраняющих его поведение. Так как каждое преобразование по объёму незначительно, то программисту легче проследить за его правильностью, а вся последовательность этих изменений может привести к существенной перестройке программы и улучшению её согласованности, четкости и простоты понимания её кода другими разработчиками.

Сведения из теории алгебраических структур

Алгебраической системой принято называть некоторое множество G (именуется носителем) с заданным на этом множестве набором различных операций и отношений (то есть сигнатурой). При этом есть частные случаи: алгебраическую систему с пустым множеством отношений называют алгеброй, а систему с пустым множеством операций — моделью.

Решёткой называют [1] множество (L, \leq) с заданным на нём отношением частичного порядка, при этом для каждой пары элементов $\{x, y\} \subseteq L$ должны существовать точные верхняя и нижняя грани ($\sup\{x, y\}$ и $\inf\{x, y\}$).

Введем, основанное на предыдущих, и понятие LP-структуры. К этому определению отнесем алгебраическую систему, которая представлена решеткой, на которой, в свою очередь, зададим некоторое бинарное отношение с логико-производционными (Logical Produce, откуда и название) свойствами. К ним относят:

- 1) рефлексивность ($\forall x \in M: (xRx)$),
- 2) транзитивность ($\forall x, y, z \in M: (xRy \cap yRz \rightarrow xRz)$),
- 3) дистрибутивность, где M -некоторое множество.

Последнее свойство подразумевает, что с помощью операций объединения и пересечения \cup и \cap , свойственных решеткам как структурам, появляется возможность логического вывода по частям и последующего объединения его результатов.

Обычно LP-структуры применяются в тех системах [2], которые подразумевают монотонный вывод, но это не всегда так. Для некоторых задач построенные модели не будут обладать этим свойством, однако, применение решеток возможно и там. Примером подобной задачи будет задача, основанная на использовании решетки типов. В качестве алгебраической решетки представим структуру типов такой объектно-ориентированной системы. Для того, чтобы задать отношение частичного порядка \leq , формализуем отношение наследования между типами, таким образом: если некоторый тип b является наследником некоторого типа a , то $b \leq a$. Тогда становится несложно определить *объединение* и *пересечение* этих двух типов. В их ролях будут выступать соответственно ближайший общий тип-предок и ближайший общий тип-потомок. Также, для придания свойства ограниченности такой решетки, введем дополнительные элементы: I – некоторый универсальный тип (предок, общий для всех типов) и O – некоторый фиктивный потомок всех типов.

Введем также и отношение R , соответствующее агрегации типов: если некоторый объект типа a содержится в типе b в качестве атрибута, то считаем $(b, a) \in R$. Это отношение, как и рассмотренное ранее (\leq и R), будут иметь общую семантику, то есть в любом из случаев, если $b \leq a$ или $(b, a) \in R$, то тип b получит возможности a при помощи доступа к его атрибутам. Также следует учесть, что общее отношение $\leftarrow (R)$ (назовем его отношением «обладания набором возможностей») должно удовлетворять свойствам рефлексивности и транзитивности. В [3] рассказывается подробнее, что это логико-производционное отношение обладает еще и ограниченным свойством дистрибутивности. Это означает, что оно будет содержать в себе заодно и семантику автоматического решения задачи «поднятия» общих атрибутов по иерархии типов, которая относится к задачам рефакторинга в объектно-ориентированном программировании. Из чего можно сделать вывод, что LP-структура, построенная в соответствии с вышеописанными условиями, дает возможность производить автоматизированные исследования иерархий типов, включая, в том числе, верификацию (использование формальных методов для доказательства правильности или неправильности программы по формальным описаниям ее свойств), эквивалентные преобразования (результат вывода программы не должен изменяться после различных

преобразований), и минимизацию (количество связей должно уменьшаться до минимального предела, не затрагивающего основной функционал программы).

UML-диаграммы программ.

UML-модель M – это представление исходного кода программы в виде совокупности некоторого множества элементов и связей. Данное множество будет конечно, так как конечен и набор связей и элементов в коде.

Рефакторингом на основе шаблона проектирования $P(r_1, r_2, \dots, r_n)$ – будем называть следующее $PR(p_1, p_2, \dots, p_s, v_1, v_2, \dots, v_m, q_1, q_2, \dots, q_k): M \rightarrow M'$ преобразование, являющееся рефакторингом, преобразующим модель M в M' , причем параметры рефакторинга связаны с ролями шаблона определенным образом. Каждому из параметров $p_i, 1 \leq i \leq s$ и $q_j, 1 \leq j \leq k$ соответствует одна из ролей $r_t, 1 \leq t \leq n$ шаблона P того же типа. Каждой роли соответствует либо один из входных параметров $p_i, 1 \leq i \leq s$, либо один из выходных параметров $q_j, 1 \leq j \leq k$ рефакторинга, либо один из входных $p_i, 1 \leq i \leq s$ и один из выходных $q_j, 1 \leq j \leq k$ параметров. Последнее возможно только в том случае, когда элементу модели соответствуют как входной, так и выходной параметры рефакторинга, т.е. этот элемент изменяется в ходе рефакторинга.

При применении рефакторинга PR к модели M в полученной модели M' обнаруживается экземпляр шаблона P , означивание которого полностью определяется фактическими значениями параметров применённого рефакторинга.

Алгоритм для использования LP-структур совместно с паттернами проектирования в рефакторинге

- 0) Составляем список шаблонов, по которым будут произведены изменения на основе LP-структур.
- 1) Создаем UML-диаграмму (модель M) по отрывку исходного кода
- 2) Анализируем ее на наличие "плохих запахов кода" (список признаков, перечисленный ранее, указывающий на признаки такого кода, которому необходим рефакторинг.)
- 3) Последовательно применяем к полученной на первом шаге модели (M) возможные рефакторинги, основанных на решетках (LP-структурах), а также набор базовых и комбинированных рефакторингов, основанных на паттернах проектирования.
- 4) В результате шагов 1-3 получаем измененную модель M'
- 5) Вычисляем для исходной и полученной моделей (M и M' соответственно) некоторые метрики, по которым можно будет оценить насколько эффективны проведенные преобразования.
- 6) Сравнение и вывод метрик и отображение пользователю изменений, для того, чтобы он также мог увидеть, что же именно произошло с исходной моделью, сравнить ее визуально с исходной и отменить часть преобразований при необходимости.

Роль рефакторинга в оптимизации программного кода

Оптимизация кода – это один из способов преобразования кода, приводящий к улучшению его характеристик и повышению производительности программы [4]. Среди целей оптимизации можно выделить уменьшение размера кода, объема используемой оперативной памяти, повышение скорости выполнения программы, уменьшение количества операций ввода – вывода. Так как под оптимизацией понимается внесение незначительных поправок, то есть изменение одного класса, одного метода или всего

лишь нескольких строк кода. Поэтому какие-либо крупные изменения проекта, приводящие к повышению производительности оптимизацией не считаются.

Понятие «оптимизация» обычно подразумевает, что система сохраняет ту же самую функциональность. Однако, значительное улучшение производительности часто может быть достигнуто и с помощью удаления избыточной функциональности. Например, если допустить, что программе не требуется поддерживать более, чем 100 элементов при вводе, то возможно использовать статическое выделение памяти вместо более медленного динамического.

Существует требование, которое обычно предъявляется к методу оптимизации – оптимизированная программа должна иметь тот же результат и побочные эффекты на тех же входных данных, что и неоптимизированная программа. Тем не менее, если изменения поведения программы, не имеет большого значения на фоне выигрыша за счет использования оптимизации, то данное требование может и не играть главной роли.

Кроме того, не существует универсального решения, которое подходило бы ко всем случаям, поэтому приходится использовать альтернативные решения, для оптимизации только ключевых параметров. Как правило, необходимые ресурсы для достижения требуемого результата, то есть получения полностью оптимальной программы, которую невозможно дальше улучшить, превышают выгоду, которую можно получить, затрачивая эти ресурсы.

Встречаются ситуации, когда оптимизированный код вручную, оказывается менее эффективнее кода, сгенерированного компилятором, поэтому особенно важно осознавать целесообразность её проведения. Оптимизация в основном фокусируется на одиночном или повторном времени выполнения, использовании памяти, дискового пространства, пропускной способности или некотором другом ресурсе. Это обычно требует компромиссов (tradeoff) – один параметр оптимизируется за счёт других. Например, увеличение размера программного кэша чего-либо улучшает производительность времени выполнения, но также увеличивает потребление памяти. Другие распространённые компромиссы включают прозрачность кода и его выразительность, почти всегда ценой деоптимизации. Сложные специализированные алгоритмы требуют больше усилий по отладке и увеличивают вероятность ошибок.

Оптимизацию производительности следует отличать от рефакторинга. Цель рефакторинга – сделать код программы более легким для понимания. Как и оптимизация, рефакторинг обычно не изменяет поведение программы. Но оптимизация часто затрудняет понимание кода, что противоположно рефакторингу.

Как правило, рефакторинг и оптимизация программного кода представляют собой разные, а иногда и вовсе взаимоисключающие процессы. Если рефакторинг, как уже говорилось выше, изменяет структуру кода для облегчения его чтения самим разработчиком, то оптимизация чаще всего производится на уже готовом коде, и улучшает не понятность исходного кода, а его скорость компиляции, быстрогодействия и состоит из замены «долгих» функций (т.е. тех, которые могут заменить собой некоторый последовательностью базовых) на те, что выполняются быстрее. Оптимизация кода, как правило, зависит еще и от используемого оборудования и технических возможностей системы. В качестве примера можно привести операцию по перемножению двух чисел. На одних системах, быстрее будет работать код, вида:

$$a = 23 * 3$$

на других же оказывается, что сложение небольшого количества чисел происходит быстрее, и код, вида:

$$b = 23 + 23 + 23$$

выполнится на доли миллисекунд раньше. Вследствие этого, глубокая оптимизация сама по себе требует обширных знаний об используемых в системе компонентах и особенностях конкретного языка программирования. Однако на практике чаще всего оказывается, что со столь глубокой оптимизацией кода на достойном уровне справляются

сами компиляторы кода. Поэтому рефакторинг, сам по себе улучшив понятность кода для программиста, может еще и оптимизировать программу просто за счет избавления от лишних или же избыточных участков этого кода, упростив её компиляцию, а также за счет возможных ошибок в тех участках кода, что были удалены, либо же переписаны на более простой, а значит и менее подверженный ошибкам, код.

Заключение

В работе показано, что рефакторинг программного обеспечения, основанный на алгебраических структурах или же на каком-либо другом методе, имеет широкий потенциал также и для применения в задачах оптимизации ПО, одновременно выполняя, на первый взгляд, противоположные задачи.

Представленный алгоритм рефакторинга моделей, описывает совместное применение групп преобразований как основанных на операциях над алгебраическими структурами, так и на паттернах проектирования, что подтверждает его достоверность и открывает широкие возможности по дальнейшему улучшению и практическому применению.

Библиографический список

1. Биркгоф Г. Теория решеток. 3-е изд., пер. с англ., М.: Наука, 1984.
2. Скорняков Л.А. Элементы теории структур. Изд. 2, пер. и доп. Букинист. 1982. 160 с.
3. Махортов С.Д. Алгебраические модели иерархий типов для проектирования и рефакторинга / С.Д. Махортов, М. Д. Шурлин // Онтология проектирования. – 2012. – № 1(3). – С. 73–79.
4. Батталова Н.И. Теоретические основы оптимизации программного кода // Материалы X Международной студенческой научной конференции «Студенческий научный форум»

L.S. Lomakina, M.N. Molchanov

REFACTORING SOFTWARE MODELS AS ONE OF THE OPTIMIZATION STEPS

Nizhny Novgorod state technical university n.a. R.E. Alekseev

The paper presents an algorithm for the refactoring algorithm using its mathematical foundations, as well as the possibility of using it with refactoring methods for program models based on design patterns and the impact of the refactoring process on program optimization.

Keywords: algebraic structures, refactoring, design patterns, UML diagrams, optimization.

НЕЙРО-НЕЧЕТКАЯ СИСТЕМА КЛАССИФИКАЦИИ СОСТОЯНИЙ ОБЪЕКТОВ СЛОЖНОЙ СТРУКТУРЫ

Нижегородский государственный технический университет им. Р.Е. Алексева

Представлен аналитический обзор по созданию нейро-нечеткой системы классификации состояний объектов сложной структуры. Сочетание искусственных нейронных сетей и систем нечеткой логики позволяет создавать устойчивые и адаптивные интеллектуальные системы. Нечеткие множества и искусственные нейронные сети являются гибкими и мощными инструментами при решении обширного спектра задач, связанных с медицинским и техническим диагностированием, а также в управлении. Проектирование комплексных систем нечеткого вывода сопряжено с необходимостью формализации достаточного количества экспертных знаний в виде ограниченного набора правил. Подготовка нейросетевой модели заключается в определении оптимальных значений для большого количества весовых параметров и требует достаточного объема данных для обучения. Искусственные нейронные сети и системы нечеткого вывода взаимно дополняют друг друга, как следствие, существует потенциал в объединении данных инструментов при решении задач классификации состояний объектов сложной структуры. В статье приведены примеры прикладных областей и аспекты эффективности нечетких множеств и нейронных сетей, а также проанализирована целесообразность применения нечетких множеств в нейросетевых классификаторах.

Ключевые слова: нечеткие множества, нейронная сеть, классификация состояний, нечеткий контроллер, нейро-нечеткий классификатор

Введение

Определение состояний объектов сложной структуры является базовой проблемой в задаче управления. В области управления повсеместно используемым подходом является применение инструмента теории нечетких множеств. Нечеткое управление считается наиболее широкой прикладной областью аппарата нечеткой логики и зарекомендовало себя как общепринятую методологию для проектирования контроллеров, обеспечивающих требуемую производительность и надежность в условиях неопределенности и неточности. Контроллеры на основе нечеткой логики могут превосходить традиционные системы управления и зачастую оказываются более эффективными, чем операторы. За последние пятьдесят лет, контроллеры на основе нечеткой логики успешно применялись для решения обширного спектра индустриальных задач [1]: управление парогенератором в лабораторных условиях, контроллер цементной печи, система очистки воды, система автоматического управления движением поездов в метрополитене города Сендай, автоматическая стиральная машина с нечетким управлением, контроллер на основе нечеткой логики также нашел применение в авиационной и космической областях. Сегодня такие контроллеры применяются по всему миру в различных областях: от бытовой техники и электроники до тяжелой промышленности (обработка нефти, цемента, изготовление стали).

Искусственные нейронные сети (ИНС) вдохновлены функционированием биологических нейронов. С развитием аппаратных возможностей исследователям удалось подтвердить теоретический потенциал ИНС и дать толчок в развитии глубокого обучения как области, в которой исследователи разрабатывают глубокие нейронные сети и методики их оптимизации. Важнейшим компонентом глубокой нейронной сети является функция активации, поскольку она играет ключевую роль в процессе обучения. С представлением миру функции активации Rectified Linear Unit (ReLU) [2], которая в настоящее время является наиболее широко используемой, исследователям и инженерам

удалось добиться значительного прогресса в разработке глубоких нейросетевых моделей для решения задач в самых различных областях, поскольку ReLU повышает эффективность обучения, в частности, благодаря противодействию эффекту «исчезающего градиента».

Нейронная сеть, как и нечеткий контроллер, является универсальным аппроксиматором в том смысле, что любая архитектура нейронной сети направлена на поиск математической функции, представляющей отображение ввода-вывода. Контроллеры аппроксимируют функцию, покрывая ее график нечеткими участками, где каждое правило в контроллере определяет нечеткий участок в пространстве ввода-вывода системы, а затем усредняет перекрывающиеся патчи.

Нечеткие множества

Инструменты аппарата теории нечетких множеств [3] характеризуются степенью неопределенности, например, классическое множество («четкое») – не обладает неопределенностью и характеризуется четким переходом между элементами (например, «холодно» и «горячо»); нечеткое множество первого типа вводит одну степень неопределенности и характеризуется «нечеткими» границами между элементами («очень холодно», «холодно», «прохладно», «нейтрально», «тепло», «горячо», «очень горячо»). Концептуально, теория нечетких множеств первого типа разрабатывалась для моделирования человеческих рассуждений, где часто встречается некоторая неопределенность, субъективные суждения. Кроме самих элементов, нечеткое множество также определяется «характеристическими» функциями, определяющими степень принадлежности элемента к множеству в промежутке от 0 до 1, где единица отражает полную принадлежность элемента нечеткому множеству, а 0 – наоборот. Таким образом, теория нечетких множеств первого типа предоставляет методологию определения промежуточных значений между значениями, связанными с абсолютной истинностью (1) или абсолютной ложностью (0). Нечеткое множество второго типа характеризуется дополнительной неопределенностью в характеристической функции (второстепенная характеристическая функция), вследствие чего принадлежность каждого элемента такого множества также описывается нечетким. Второстепенная характеристическая функция добавляет новое изменение в пространство области определения нечеткого множества первого типа. Нечеткое интервальное множество второго типа отличается от обозначенного ранее аналога тем, что накладывает ограничение на форму второстепенной характеристической функции – у нечеткого множества второго типа форма может принимать вид любой функции из определенных в множестве первого типа (треугольную, форму трапеции или сигмоиды), а у интервального множества только прямоугольную.

Многочисленные исследования показывают, что контроллеры, в основе которых заложен аппарат теории нечетких множеств 2-го типа, превосходят аналогичные контроллеры 1-го типа, а также, многие традиционные контроллеры. Это во многом объясняется тем, что использование нечетких множеств 2-го типа снижает базу правил контроллера, такие контроллеры имеют более гладкую поверхность функции управления и поддерживают более комплексные отношения ввода-вывода. Все это теория нечетких множеств объясняет тем, что нечеткое множество 2-го типа содержит в себе набор нечетких множеств 1-го типа. На практике перечисленные преимущества приводят к стабильной работе контроллера, работающего на теории нечетких множеств 2-го типа в условиях, при которых другие контроллеры перестают работать стабильно, например, управление скоростью морских дизельных двигателей, управление уровнем жидкости в спаренном резервуарном аппарате, управление температурой горячей прокатки стали, управление станком для микрофрезеровки, управление высотой полета самолета, управление навигацией мобильных роботов.

Нейронные сети и нечеткие множества

Сочетание искусственных нейронных сетей и систем нечеткой логики дает возможность решать задачи путем создания устойчивых, интеллектуальных и адаптивных систем. В ряде исследований [1] показывается, что нечеткие логические системы (FLSs) являются гибкими инструментами для обработки неопределенности и определения нелинейных отображений ввода–вывода.

Потенциал применения нейро-нечетких алгоритмов в области медицины, сосредотачивается на методах, в которых нейронные сети в сочетании с инструментом теории нечетких множеств могут быть использованы для повышения эффективности распознавания закономерностей. Несмотря на то, что медицина и инженерия управления не обладают прямой взаимосвязью, использование методов управления, для доступных онлайн устройств, особенно в случаях хирургических операций и в отделениях интенсивной терапии, в настоящее время уже возможно. Сегодня прикладные области технологии управления в медицине варьируются от простых систем назначения дозировок лекарств до высокоразвитых адаптивных контроллеров. Исследования в [4] показали, что методы нечеткой логики могут быть использованы в ранней диагностике таких заболеваний, как болезнь Паркинсона, для выявления рака груди.

Показательным примером является проблема диагностики биоценоза. Биоценоз – совокупность разных видов микроорганизмов, связанных между собой определенными отношениями и населяющих определенную биологическую нишу. Частный случай биоценоза – микробиота желудочно-кишечного тракта (ЖКТ) человека. Это количественно-качественный состав его микрофлоры, то есть заселяющих его микроорганизмов. Биоценоз является сложным медико-биологическим объектом, для классификации которого классические методы обработки данных имеют недостаточную эффективность.

В [5] на основе метода проекций проведен анализ данных бактериологического исследования микрофлоры кишечника, предоставленных Нижегородским научно-исследовательским институтом эпидемиологии и микробиологии им. Академика И. Н. Блохиной. В каждом из классов (болен, здоров) по экспериментальным данным вычисляются оценки математического ожидания и дисперсии проекции случайного вектора ξ . Пороговое значение, которое делит пространство на критическую и допустимую области (рис. 1) при равенстве априорных вероятностей того здоров пациент или болен вычисляется при $L=1$ и по формуле [1].

$$x_0 = \frac{\sigma_2^2 M_1 - \sigma_1^2 M_2 \pm \sigma_1 \sigma_2 \sqrt{(M_2 - M_1)^2 + (\sigma_2^2 - \sigma_1^2) \ln \frac{\sigma_2^2}{\sigma_1^2}}}{\sigma_2^2 - \sigma_1^2}, \quad (1)$$

где L – критерий отношения правдоподобия, M_1 и M_2 – математические ожидания проекций групп здоровых и больных пациентов, и σ_1 и σ_2 – среднеквадратические отклонения.

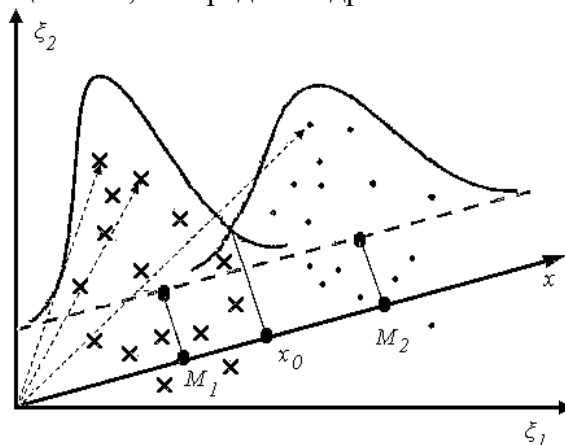


Рис. 1. Визуализация метода проекций

На основании расположения проекции x относительно заранее вычисленного значения величины принимается решение о принадлежности к классу. Если значение проекции окажется близким к значению порога, то классификация с использованием данного метода будет иметь низкую точность. Для решения задачи классификации в [6] предлагается использовать ИНС, в частности сеть на радиально-базисных функциях (RBFN). ИНС подбирает скрытые внутренние параметры как отдельных нейронов, так и всей сети в целом, а также выделяет наиболее актуальные входные признаки и отбрасывает малоинформативные, некорректные признаки. Эта особенность позволяет ИНС исправно работать с большим количеством неинформативных и избыточных данных.

В последнее время исследователи начали уделять больше внимание интервальным нечетким логическим системам 2-го типа (IT2-FLS), поскольку они лучше справляются с неопределенностями и способны определять нелинейные отображения ввода-вывода, которые не могут быть достигнуты нечеткими логическими системами 1-го типа (FLS). Таким образом, внедрение IT2-FLSs [7] стало популярным в приложениях с высокой неопределенностью и сложностью, таких как проектирование систем управления, предсказание временных рядов, вычисления со словами, принятие решений. IT2-FLS используют и применяют интервальные нечеткие множества 2-го типа (IT2-FS), которые имеют дополнительную степень свободы, обеспечиваемую отпечатком неопределенности (FOU), что является источником их способности справляться с неопределенностями. Внутренняя структура IT2-FLS схожа с аналогичной структурой первого типа с той лишь разницей, что в ней присутствует механизм сокращения типа (TR), поскольку она построена на нечетких множествах второго типа. Несмотря на то, что IT2-FLS эффективны для моделирования неопределенностей и нелинейностей из-за дополнительной степени свободы, обеспечиваемой их отпечатком неопределенности, они также обладают внушительным количеством настраиваемых параметров, усложняющих проектирование. Чтобы преодолеть это узкое место, нейронные сети были использованы для решения проблемы проектирования IT2-FLS, поскольку они обладают превосходной способностью к обучению.

Нейро-нечеткий классификатор

Ярким примером нейро-нечеткой системы является сеть на основе адаптивной нейро-нечеткой системы вывода (ANFIS). ANFIS принадлежит к семейству гибридных нейро-нечетких систем и совмещает в своей структуре свойства нейронных сетей и принципы нечеткой логики. Эта структура использует методы машинного обучения для определения параметров системы нечеткого вывода на основе IF-THEN правил для описания поведения сложных систем.

Структурная схема модели ANFIS представлена на рис. 2. Модель реализует контроллер с несколькими входами и правилами.

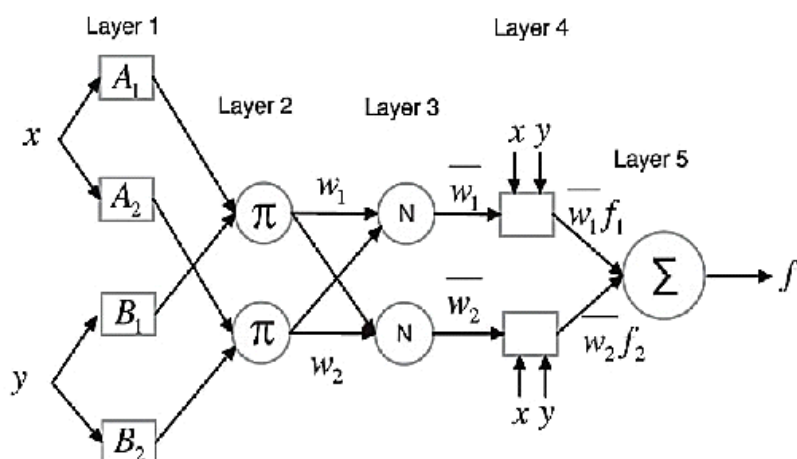


Рис. 2. Архитектура адаптивной нейро-нечеткой системы вывода ANFIS.

В [7] рассматривается применение теории нечетких интервальных множеств второго типа в алгоритмическом обеспечении глубокой нейросетевой модели в качестве функции активации Interval Type 2 Fuzzy Rectified Unit (IT2FRU). Авторы представили «нечеткий слой активации» (FAL), улучшающий эффективность глубоких нейронных сетей. Представленная функция активации IT2FRU обладает тремя конструктивными параметрами, которые гибко задают форму отображения ввода–вывода. Гибкость в данном контексте означает возможность изменения размера отпечатка неопределенности нечетких интервальных множеств второго типа, благодаря чему процесс обучения глубокой нейронной сети происходит заметно эффективнее. Также, функция активации способствует снижению проблемы «исчезающего» градиента и ускорению сходимости модели глубокой ИНС. Другой отличительной особенностью такой нейронной сети является стабильность ее функционирования при различных конфигурациях гиперпараметров процесса обучения.

ИНС на нечетких слоях активации «уверенно» справляется с классическими задачами компьютерного зрения – распознаванием рукописных цифр, классификацией изображений и, по полученным результатам, превосходит аналогичные модели ИНС с обычными функциями активации.

Результаты и выводы

В статье проанализировано применение теории нечетких множеств для классификации состояний сложных объектов в теории управления. Проведенный обзор преимуществ и областей применения указывает на перспективу дальнейшего развития данного инструментария.

Применение теории нечетких множеств 2-го типа в нейросетевых классификаторах состояний объектов сложной структуры в нейро-нечетких классификаторах позволит стабилизировать процесс обучения, снизить время самого обучения, увеличить точность классификации, а также эффективнее решать проблему переобучения. Обозначенные аспекты свидетельствуют о потенциале описанного подхода в задаче классификации состояний объектов сложной структуры.

Библиографический список

1. Mendel J., Nagraj H., Woei Wan Tan, Melek W. and Ying H. "Introduction to Type-2 Fuzzy Logic Control: Theory and Applications." 2014.
2. Nair, Vinod, and Geoffrey E. Hinton. "Rectified linear units improve restricted boltzmann machines." Icm1. 2010.
3. Кофман А. Введение в теорию нечетких множеств [Текст] / А. Кофман / пер с франц. – М.: Радио и связь, 2007. – 432 с
4. Коробова Л.А., Гладких Т.В. Разработка модели принятия решения для постановки диагноза заболеваний на основе нечеткой логики. Вестник Воронежского государственного университета инженерных технологий. 2018;80(4):80-89
5. Ломакин Д.В., Ломакина Л.С., Пожидаева А.С. Вероятность. Информация. Классификация: учеб. пособие. Нижегород. гос. техн. ун–т им. Р.Е. Алексеева. – Н. Новгород, 2013. – 134 с.
6. Ломакина Л.С., Носков К.М. Нейросетевые технологии диагностирования состояний биоценоза на основе априорных статистических данных // Труды НГТУ №1, 2018. 37-43
7. Aykut Beke, Tufan Kumbasar Learning with Type-2 Fuzzy activation functions to improve the performance of Deep Neural Networks. Engineering Applications of Artificial Intelligence

NEURO-FUZZY CLASSIFICATION SYSTEM FOR CLASSIFICATION OF THE STATES OF OBJECTS OF COMPLEX STRUCTURE

Nizhny Novgorod state technical university n.a. R.E. Alekseev

Purpose: Fuzzy sets and artificial neural networks are flexible and powerful tools for solving a wide range of tasks related to medical and technical diagnostics, as well as in control. The article considers the combination of fuzzy sets theory and neural network theory to increase the efficiency of deep learning models and provides examples of application areas and aspects of the effectiveness of fuzzy sets and neural networks, as well as analyzes the feasibility of using fuzzy sets in neural network classifiers.

Design/methodology/approach: system analysis, fuzzy set theory, neural network theory.

Findings: The combination of artificial neural networks and fuzzy logic systems allows to create stable and adaptive intelligent systems.

Research limitations/implications: The design of complex fuzzy inference systems involves necessity to formalize a sufficient amount of expert knowledge in the form of a limited set of rules. The preparation of a neural network model consists in determining the optimal values for a large number of weight parameters and requires a sufficient amount of data for training.

Originality/value: The article provides examples of applied areas and aspects of the effectiveness of fuzzy sets and neural networks, as well as considers the feasibility of using neuro-fuzzy classification systems in a number of tasks that require determining the state of objects of complex structure.

Keywords: fuzzy sets, neural network, state classification, fuzzy controller, neuro-fuzzy classifier.

ИССЛЕДОВАНИЕ ПРОИЗВОДИТЕЛЬНОСТИ РАСПРЕДЕЛЕННОГО ПРОГРАММНОГО ГЕНЕРАТОРА ГРАФ-МОДЕЛЕЙ

Нижегородский государственный технический университет им. Р.Е. Алексеева

Целью работы является создание среды для написания генераторов графовых моделей большой размерности с использованием параллельного программирования. Автором создана программная среда, в которой благодаря спецификациям большая часть работы по написанию кода окружения генератора выполняется с помощью простого конфигурирования системы.

Ключевые слова: генератор, графовые модели, графы большой размерности, параллельное программирование, мультиагентная среда.

Введение

При разработке различных алгоритмов, систем, нейронных сетей и т. д. входные данные нередко представляются в виде графовых моделей. Для тестирования этих проектов нужны графы различных типов и конфигураций. Для создания тестовых данных создают генераторы графов. Однако для качественного тестирования размер входных данных должен быть достаточно велик. Одной отдельно взятой машине сгенерировать такой объём невозможно, поэтому применяют параллельное программирование для генерации графов на кластерах.

Однако для работы с генераторами и полученными данными нужна определенная среда, которая сможет обрабатывать полученные данные, переносить их с носителя на носитель и т. д. Параллельное программирование достаточно трудоемко и также требует организации определенной среды в соответствии с выбранной технологией.

Поэтому создание многофункциональной универсальной среды, в которой основную часть работы по созданию окружения генератора система берет на себя, очень актуально и позволит упростить весь процесс. Созданная автоматизированная система должна просто конфигурироваться за счет спецификаций. В идеале данная среда должна разгрузить разработчика генераторов, при этом поддерживая и помогая ему работать непосредственно с алгоритмами генерации, не отвлекаясь на организацию периферийных задач.

Существующие аналоги

Найти фреймворк, который рассчитан на распределенную генерацию графов, не удалось. Все фреймворки предполагают либо только простую работу с графами, либо рассчитаны лишь на распределенные вычисления, а логику графов придется встраивать в них вручную. Поэтому единственный вариант: объединить 2 фреймворка, что уже не очень удобно, так как требует интеграции двух совершенно разных сред. При этом данные фреймворки будут содержать массу избыточной функциональности, которая затруднит их изучение.

Перечислим фреймворки, рассчитанные на работу с графами:

1. JUNG — фреймворк с открытым исходным кодом для моделирования графов и их визуализации. Фреймворк имеет ряд встроенных алгоритмов компоновки, а также алгоритмы анализа, такие как кластеризация графов и метрики для центральности узлов. Достаточно неплохой фреймворк, однако обладает довольно запутанной документацией и последние несколько лет не обновляется, поэтому его можно считать устаревшим.

2. Guava common. graph (GCG) — библиотека, входящая в популярный фреймворк Guava. Имеет простую и понятную структуру, есть немало полезных классов и абстракций, облегчающих работу с графами. Однако все окружение рассчитано

исключительно на работу с объектами внутри приложения, переносом графа из приложения на постоянные носители придется заниматься самостоятельно. Также стоит понимать, что данная библиотека входит в набор Guava, что накладывает определенные ограничения на неё. Библиотека развивается и постепенно начинает интегрироваться с более специализированными библиотеками, однако на данный момент обладает недостаточным функционалом для простой работы с графами.

3. JGraphT — программный компонент с открытым исходным кодом для рисования графов, доступный для языков Java, C # и JavaScript. Имеет массу встроенных методов для создания и всевозможной обработки графов, а также множество шаблонов, базируясь на которых можно легко сгенерировать любую желаемую топологию графа. Есть возможности для гибкой настройки импорта и экспорта графов, визуализации и т. д. Перечислим средства для распределенных вычислений:

OpenMP — открытый стандарт для распараллеливания программ на языках Си, Си++ и Fortran. Дает описание совокупности директив компилятора, библиотечных процедур и переменных окружения, которые предназначены для программирования многопоточных приложений на многопроцессорных системах с общей памятью. OpenMP не подходит для данной задачи, так как рассчитан на многопроцессорные системы с общей памятью, но не для кластеров.

MPI — программный интерфейс (API) для передачи информации, который позволяет обмениваться сообщениями между процессами, выполняющими одну задачу. Однако данный интерфейс подразумевает довольно низкоуровневое программирование, чрезвычайно усложняет код, а также слабо поддерживается сообществом, поэтому он не упрощает разработку распределенных программ, а лишь позволяет контролировать взаимодействие между ними на максимально базовом уровне.

JADE — программная среда разработки мультиагентных систем и приложений, поддерживающая FIPA-стандарты для интеллектуальных агентов. Довольно развитый фреймворк, агенты в данном фреймворке достаточно универсальны, однако из-за этого довольно много весят, что отрицательно сказывается на производительности в больших системах. Также агенты являются «активными слушателями», то есть постоянно сканируют внешнюю среду, тратя на это много ресурсов системы. Эти особенности фреймворка не позволяют эффективно использовать ресурсы кластера, так как затраты на обслуживание фреймворка линейно возрастают с увеличением масштабов системы.

Akka — фреймворк для создания параллельных, распределенных и отказоустойчивых приложений, управляемых сообщениями. Akka реализует парадигму акторов. Данная парадигма схожа с парадигмой агентов. Акторы достаточно легковесны и активируются только при получении сообщения, которые обрабатывают диспетчеры, поэтому не потребляют ресурсы в режиме ожидания. Данные особенности делают фреймворк эффективным для распределенных приложений, так как система характеризуется отличной масштабируемостью. При этом разработчик ПО на основе данного фреймворка должен лишь сконфигурировать акторную систему и разработать систему сообщений между акторами, всю остальную работу по обеспечению работоспособности системы со всеми присущими ей свойствами берет на себя фреймворк.

Некоторые фреймворки отлично справляются со своей работой, однако ни один из них не поддерживает напрямую обе возможности, которые нам нужны. Для получения необходимого результата нужно скомбинировать два фреймворка, что является достаточно непростой задачей и требует длительного изучения обоих фреймворков, при этом настаивать нужно оба. Довольно сложно обеспечить грамотное взаимодействие столь разных по назначению решений. Именно поэтому и было решено создать собственное, которое бы позволяло абстрагироваться от настроек данной интеграции и обойтись простым конфигурированием в зависимости от задачи. К тому же было необходимо обеспечить хорошую масштабируемость проекта, так как вариантов задач, которые решает данный фреймворк, очень много.

Описание системы

За основу были взяты фреймворки JGraphT и Akka. JGraphT обеспечивает работу с графами. На его основе выстраивается логика создания, загрузки и сохранения графа. Система поддерживает ввод и вывод графа в файл и БД различных типов, а также вывод в консоль.

За работу с БД отвечает специальный JDBC-контроллер, который предоставляется в виде менеджера. Данный контроллер предоставляет все необходимые в работе методы, основанные на SQL-запросах в данные БД, в том числе и некоторые оптимизационные, сокращающие количество транзакций до оптимального минимума (так как каждая транзакция тратит много ресурсов). Контроллер поддерживает базы данных MySQL и PostgreSQL, однако при необходимости несложно добавить поддержку прочих БД.

За распараллеливание отвечает Фреймворк Akka. Есть главный актер — MainProcess, который порождает рабочих акторов — Process-ы. Также MainProcess распределяет задачу между Process-ами и выполняет ту часть генерации, которая не может быть проведена в параллельном режиме. Для того, чтобы создать свою версию генератора, нужно взять за основу уже написанные MainProcess и Process, и реализовать определенные для этого методы. Вся остальная логика, в том числе и сбор статистики, уже заложена в данные классы.

Также реализован пользовательский интерфейс, который использует REST-API, позволяющее в браузере выбрать необходимую конфигурацию по работе с графом. Это загрузка или сохранение графа, генерация графа с нуля, продолжение генерации в уже существующий граф. Все легко масштабируется, новые генераторы в систему добавляются парой простых манипуляций.

Разработка новых генераторов происходит достаточно быстро и эффективно, количество рутинного кода сведено к минимуму. Все, что необходимо разработчику для создания генератора — следовать спецификации и инструкции и правильно настраивать систему. Среда создает слой абстракции, который отгораживает разработчика от тонкостей работы распределенной среды, позволяет по максимуму писать код как для однопоточного приложения. Также среда предоставляет «из коробки» все необходимые инструменты для обработки графа и взаимодействия с БД.

Исследование производительности фреймворков.

Был проведен эксперимент по производительности генератора в однопоточном режиме, режиме распараллеливания с помощью MPI, и режиме распараллеливания средствами Akka.

Тестирование проводилось на одном и том же компьютере с двухъядерным процессором, генерировался один и тот же граф одних и тех же размеров. Каждый опыт проводился несколько раз для исключения влияния загруженности системы и оценки устойчивости системы при каждом конкретном наборе входных условий.

Однопоточному режиму соответствует строка с количеством потоков = 1. Данное сравнение позволяет оценить накладные расходы на работу системы, оно дает понять, как много времени уходит на стартовую инициализацию и сравнить обслуживание какого фреймворка обходится дороже в плане ресурсов.

Для MPI запустить систему с количеством потоков больше 20, не удалось, так как стартовая инициализация каждого потока в MPI создает большую нагрузку на ОЗУ компьютера, и в итоге при большом количестве экземпляров потоков происходит вылет из-за переполнения heap виртуальной машины JVM. В Akka акторы очень легковесны (порядка 300 Кб) и создаются по требованию, поэтому подобных проблем не наблюдается. По этой же причине, однопоточная программа на Akka выполняется гораздо быстрее, так как стартовая инициализация происходит довольно быстро.

Сравнение производительности MPI и Akka

Количество потоков	Время генерации MPI, с	Время генерации Akka, с
1	14,4	10,9
2	11,6	9,9
3	10,9	9,1
4	11,0	8,0
5	12,2	9,9
6	12,0	7,0
7	12,5	9,6
8	12,5	9,8
16	11,6	9,4
20	17,7	6,2
32	-	10,2
100	-	9,7
320	-	9,4
640	-	9,3
1000	-	10,8
10000	-	12,5
100000	-	15,3

MPI показала сильную зависимость времени генерации от количества потоков, идеальным оказалось 3 потока на 2 ядра, то есть 2 потока занимают ядро, а 3-ий поток перехватывает свободное процессорное время, которое периодически возникает из-за различных задержек при работе потоков.

Akka куда менее зависима от количества потоков, так как она лучше распределяет нагрузку, идеальным оказался показатель 20 потоков, однако разница во времени с остальными вариантами не столь заметна и на больших объемах задач с большими ресурсами станет пренебрежимо малой. При этом даже на столь слабом компьютере Akka эффективно справилась даже со 100000 потоками без существенных потерь производительности.

По всем показателям Akka превзошла MPI. По сравнению с однопоточным режимом выгода не столь очевидна, но она есть. На больших мощностях данная разница будет гораздо заметнее, так как в данном случае будут подключаться не только дополнительные виртуальные потоки, но и реальные ядра. Выигрыш на одном процессоре обуславливается оптимизацией использования процессорного времени, которое итак неплохо оптимизируется средствами JVM, однако даже в этом случае выигрыш использования Akka есть.

Заключение

В ходе работы была создана среда для эффективного написания генераторов графов большой размерности с использованием параллельного программирования. В единую систему были объединены фреймворки JGraphT и Akka, создано окружение для работы с графами. Система была протестирована на производительность и эффективность, а также легкость написания кода.

Библиографический список

1. Официальный сайт фреймворка JUNG // [Режим доступа]: <https://jrtom.github.io/jung> (Дата обращения: 27.01.2021)
2. Документация библиотеки GCG // [Режим доступа]: <https://guava.dev/releases/20.0/api/docs/com/google/common/graph/package-summary.html> (Дата обращения: 27.01.2021)
3. Официальный сайт библиотеки JGraphT // [Режим доступа]: <https://jgrapht.org> (Дата обращения: 27.01.2021)
4. Официальный сайт API OpenMP // [Режим доступа]: <https://www.openmp.org> (Дата обращения: 27.01.2021)
5. Официальный сайт библиотеки MPI-Express // [Режим доступа]: <http://www.mpj-express.org> (Дата обращения: 27.01.2021)
6. Официальный сайт фреймворка JADE // [Режим доступа]: <https://jade.tilab.com> (Дата обращения: 27.01.2021)
7. Официальный сайт фреймворка Akka // [Режим доступа]: <https://akka.io> (Дата обращения: 27.01.2021)

A.S. Rassadin

EXPERIMENTAL STUDY OF A DISTRIBUTED SOFTWARE GRAPH MODEL GENERATOR

Nizhny Novgorod state technical university n.a. R.E. Alekseev

Purpose: creation of an environment for writing generators of large-scale graph models using parallel programming.

Design/methodology/approach: parallel programming, graph theory

Findings: an environment was created for efficiently writing generators of high-dimensional graphs using parallel programming.

Research limitations/implications: the environment can be used to write graph generators of various types using parallel programming

Originality/value: the environment can be used to generate test data for neural networks and search algorithms

Keywords: generator, graph models, high-dimensional graphs, parallel programming, multi-agent environment.

АЛГОРИТМ ГЕНЕРАЦИИ МУЗЫКИ НА ОСНОВЕ АВС-НОТАЦИИ И ДИСТРИБУТИВНОЙ СЕМАНТИКИ

Нижегородский государственный технический университет им. Р.Е. Алексеева

В данной работе предложен алгоритм генерации музыки, который основан на представлении музыкальных произведений в АВС-нотации, используя модель дистрибутивной семантики Word2Vec, современную архитектуру нейронных сетей – Трансформер. Полученные результаты показали, что алгоритм генерирует музыку не хуже, чем обычное случайное микширование заготовленных музыкальных фрагментов, получаются осмысленные мелодические и ритмические конструкции.

Ключевые слова: генерация музыки, АВС-нотация, модель word2vec, нейронная сеть transformer, конечные автоматы.

Введение

Искусственный интеллект, благодаря развитию вычислительных систем очень улучшился за последнее десятилетие. За это время было разработано огромное количество архитектур нейронных сетей, которые смогли качественно обрабатывать даже самые сложные данные: изображения, тексты на естественном языке и прочее. Однако, не так много работ посвящено задачам обработки музыки, особенно генерации новых произведений, хотя это является одной из ключевых теоретических и практических проблем в области искусственного интеллекта. Поэтому была выбрана именно эта область исследования, целью которой является разработка алгоритма генерации музыки.

Основные подходы к генерации музыки

Основных подходов к генерации музыки всего три – генерация на основе заранее созданных фрагментов и дальнейшее случайное микширование их, генерация на основе нотной записи и генерация музыки из «сырого» звука. Самый распространенный метод – первый, им пользуются не первое десятилетие и в целом данный подход показывает себя хорошо, но нельзя его назвать «интеллектуальным», так как большую часть работы делает человек. Третий подход – самый перспективный в развитии, но при этом один из самых сложных, так как необходимо решить огромный пласт задач обработки сырых сигналов для извлечения нужных признаков, чтобы непосредственно перейти к задаче генерации музыки. Второй подход на данный момент является самым актуальным, так как в области обработки естественного языка в 2018 году произошел прорыв, благодаря модели ELMo, архитектуре BERT от Google нейронной сети transformer [1], именно поэтому в данном докладе предлагается сведение работы с музыкой к работе с текстом, используя современные архитектуры нейронных сетей и дистрибутивную семантику.

АВС-нотация

Для того, чтобы эффективно использовать нотную грамматику в вычислительных системах был разработан полноценный язык разметки АВС в 1994 году [2]. Именно в таком формате хранятся ноты в электронных носителях информации. АВС-нотация состоит из трех частей: первая часть – заголовочная, где хранится общая информация: размерность, тональность, композитор, наименование, темп и прочее. Вторая часть – основная, где описываются ноты, ритмы, различные конструкции, которые показывают в нотах акценты, динамику, стиль игры. В третьей части прописываются слова к произведению, если они есть. На рис.1 представлено как выглядит пример небольшого музыкального фрагмента на АВС-нотации и в классической нотации.

В рамках исследовательской работы был использован датасет, который предоставляла компания Yandex.Cloud. В датасете 190 тысяч произведений в формате ABC-нотаций различных мелодий многих жанров, даже которые использовались в компьютерных играх для поддержания атмосферы. А задача ставится следующем образом: разработать алгоритм, который на вход получает восемь тактов музыкального произведения, алгоритм должен предоставить продолжение произведения на следующие 8 тактов. Для решения задачи в данной работе предложено использовать современные архитектуры нейронных сетей и дистрибутивную семантику.

Softpedia

```
X: 1
T: Softpedia
M: 3/4
L: 1/16
K: B
=A, g[E,B]^E,2^^f2 ^F2 F,=E, Fb2F,2 | F,5^e4^^E2^f4e4E3 |g2 g2[=A5^^E5:
B4 e3 ^^F7 |^B4=B7 |a2b2a8 |^e=E4^^F,4F4de2E2 |
f2F,5^E,2E,=e2e4 |^e3B |
```

Рис. 1. ABC-нотация

Дистрибутивная семантика

Дистрибутивная семантика – это область лингвистики, которая занимается вычислением степени семантической близости между лингвистическими единицами на основании их распределения в больших массивах текстов. Каждой лингвистической единице присваивается конкретный вектор, а множество таких векторов формирует векторное пространство. Зачастую, под лингвистической единицей подразумевают полноценные слова или n-граммы.

Современными моделями для дистрибутивной семантики являются Word2Vec, fastText, GloVe, ELMo и прочее. Выбор модели — отдельное пространство для практических экспериментов. В рамках первоначального исследования будет представлено считающуюся уже классической моделью для векторизации — Word2Vec.

Word2Vec – это трехслойная нейронная сеть, существуют две ее основных реализаций– Skip-gram и CBoW. CBoW — архитектура нейронной сети, которая предсказывает текущее слово, анализируя окружающий его контекст. Архитектура Skip-gram использует текущее слово, чтобы предсказать окружающие его слова [3].

Архитектура нейронной сети Skip-gram состоит из следующих слоев:

1. Входной слой: принимает вектор слов в one-hot encoding, то есть слово помечается единицей в словаре, а длина one-hot вектора равняется размеру словаря.

2. Слой Embedding, который представляет собой матрицу размера $N \times P$, где N – размер входного вектора (размер словаря), а P – это величина векторного пространства дистрибутивной семантики, который подбирается эмпирически.

3. Выходной слой Softmax, размерностью $N \times 1$. Каждый из нейронов в этом слое выдает вероятность того, что входное слово принадлежит соответствующему контексту.

Архитектура «непрерывного мешка слов» - полностью зеркальна Skip-gram, то есть на вход подается контекст, а сеть предсказывает слово, подходящее контексту.

На основе подаваемых текстов и обучается данная нейронная сеть, которая рассматривает соседние слова, как контекст. Но после обучения самое главное для нас получить те самые вектора слов – embedding'и. И это сделать достаточно просто: необходимо умножить one-hot вектор на матрицу embedding, тем самым получив значение вектора определенного слова. Данные вектора слов будут использоваться в основной нейронной сети для генерации мелодических и ритмических конструкций.

Нейронная сеть Transformer

Компания Google в 2017 году представил новую архитектуру нейронных сетей – Трансформеры. И после этого открытия начали развивать данную архитектуру и предложили большое количество архитектур: GPT, T5, Bert, ALBERT, и все они относятся к семейству трансформеров. Данные нейронные сети сделали огромный рывок в области NLP, именно поэтому пал выбор именно на них. Однако, тут также большое поле для экспериментов выбора архитектур.

Для задач генерации текста, таких как автоматический перевод, аннотирование и прочего, была разработана модель Энкодера-Декодера, где зачастую использовались рекуррентные нейронные сети. Однако, рекуррентные нейронные сети слабо распараллеливаются, а, следовательно, тратится огромное количество времени на обучение и использование [4]. Поэтому и была разработана модель трансформера с механизмом внимания. На рисунке 2 представлена архитектура из статьи Google [5].

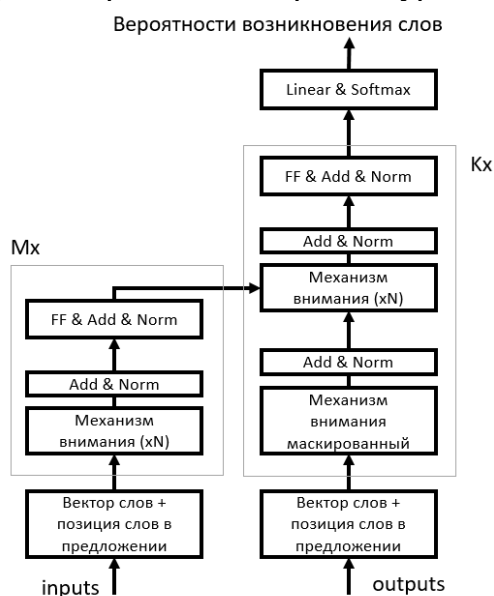


Рис. 7. Архитектура Трансформер.

Трансформер работает по схеме энкодер-декодер. Рассмотрим, как работает энкодер:

- 1) Производится векторизация каждого слова в предложении, но дополнительно добавляется информация о расположении слова в нем.
- 2) Для первого слова вычисляется «степень похожести» со всеми другими.
- 3) «Степень похожести» умножается на вектор второго слова, тем самым получаем некоторую концентрацию уже двух слов.
- 4) После такой обработки всех слов, полученные концентрации складываются друг с другом, получая один результирующий вектор предложения. Именно это и является «механизмом внимания». Однако, считается недостаточным использовать лишь один механизм, так что, как правило, используется несколько таких механизмов, чтобы находить скрытую взаимосвязь.
- 5) Результирующие векторы конкатенируются, а после чего производится снижение размерностей до исходной размерности вектора. В результате имеем вектор предложения с механизмами внимания относительно первого слова.
- 6) Далее получаем такие же вектора для всех остальных слов.

Так как полученные вектора имеют такую же размерность, что и входные, то можно использовать блок п.2-п.6 несколько раз, но уже не с векторизованными словами, а тем, что получается на выходе. И в результате получается сложный embedding для каждого слова, который и будет обрабатывать декодер.

Рассмотрим алгоритм работы декодера:

Декодер запускается последовательно по одному слову. И должен выдать следующее. Однако тут появляется два вида использования «механизма внимания»:

- 1) возможность обратиться к векторам прошлых декодированных слов, также, как и было в процессе энкодинга.
- 2) Возможность обратиться к выходу энкодера. В этом случае на механизм внимания подаются вектора с энкодера, а также вектора с информацией, полученной декодером ранее.

В конце нейронной сети стоит слой softmax, который выдает вероятности появления того или иного слова [6].

Описание алгоритма

Алгоритм построен на 4 нейронных сетях: две модели word2vec и два transformer'a. Каждую из нейронных сетей необходимо предварительно обучить, так как ритмические и мелодические конструкции не сводятся к обычным словам, а являются наборами символов. Используется датасет, представленный Yandex.Cloud, он состоит из 190 тысяч музыкальных произведений различных жанров в формате ABC-нотации.

Предложенный алгоритм генерации музыки состоит из следующих этапов:

На первом этапе с помощью детерминированного конечного автомата (далее К.А.) производится разделение ABC-нотации на мелодические и ритмические конструкции. В данном исследовании используется следующая специфика: один аккорд – одно слово. Если написана одна нота – то это тоже отдельное слово. Один такт – одно предложение.

На втором этапе для каждого текста конструкций производится векторизация с помощью word2vec, которую необходимо изначально обучить. Обучение происходит на уже разделенных текстах конечным автоматом. В итоге получаем две модели word2vec, которые будут использованы для векторизации. Если при векторизации новой ритмической конструкции или аккорда слово не найдено и не было использовано до этого, то берется самая нижняя нота аккорда, как образующая, и используется вектор для нее, как характерный для аккорда.

На третьем этапе с помощью нейронных сетей transformer производится генерацию продолжения мелодических и ритмических конструкций соответственно. Также необходимо предварительно обучить на данных, с использованием векторов слов после word2vec.

На четвертом этапе с помощью конечного детерминированного автомата сцепляем мелодию и ритм для получения музыкального продолжения в формате ABC-нотации, которую возможно перевести в музыкальный формат и прослушать получившееся музыкальное продолжение.

На рис.3 представлена схема алгоритма.

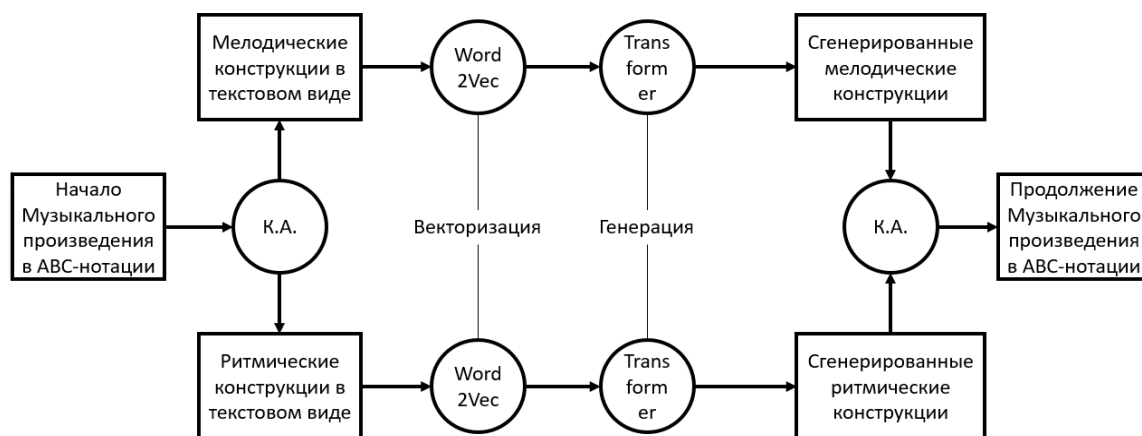


Рис. 3. Схема предложенного алгоритма

На данный момент успешно реализовал конечный автомат разделения, а также обучена модель word2vec для дистрибутивной семантики. И так как word2vec способен генерировать продолжение, смотря на контекст, был опробован метод генерации на основе word2vec. Однако, результаты не устраивают и на данный момент обучаются трансформеры для более качественной генерации, так, как и модель сложнее и механизмы лучше себя показывают в задачах генерации текстов, а модель word2vec небольшая, и основная ее цель – дистрибутивная семантика.

Заключение

В данной работе был представлен алгоритм генерации музыки, а также описаны использованные модели для достижения генерации качественной музыки. Задача по праву является очень интересной и мало раскрытой на сегодняшний день. В дальнейшем планируется полностью реализовать предложенный алгоритм, подобрать параметры для эффективной работы. Но стоит сказать и о возникающих сложностях, связанных с оценкой качества получаемых музыкальных произведений? Ведь нельзя просто сравнивать сгенерированную музыку машиной и написанную композитором, ведь возможно, машина сможет сделать даже лучше. Поэтому, задача выбора метрики настолько же сложна, как и решение непосредственно самой задачи. В будущем, планируется оценивать качество алгоритма на основе опроса людей, где изначально произведение и сгенерированное будет засекречено, а люди будут голосовать за более понравившийся музыкальный фрагмент. Далее, полученные результаты будут подвергнуты статистическому анализу и приведены в некоторое цифровое значение. Тогда, на основе сравнения будет явно понятно на сколько хорошо алгоритм смог сгенерировать продолжение у музыкального произведения.

Библиографический список

1. BERT, ELMO и Co в картинках (как в NLP пришло трансферное обучение) / JayAlamar, Смирнова Екатерина, Шкарин Сергей - <https://habr.com/ru/post/487358/>
2. A brief history of abc / Chris Walshaw - <http://abcnotation.com/history>
3. Distributed Representations of Words and Phrases and their Compositionality / Tomas Mikolov, Ilya Sutskever, Kai Chen, Greg Corrado, Jeffrey Dean / Google Inc, 2013 – 9 стр.
4. Суркова А.С. Использование модели Encoder-Decoder для реферирования текстов / Суркова А.С., Чернобаев И.Д. / Системы управления и информационные технологии, №4(70), 2017. – С. 72-76.
5. Attention Is All You Need / Ashish Vaswani, Noam Shazeer, Niki Parmar, Jakob Uszkoreit, Llion Jones / Google Inc, 2017 – 15 стр.
6. Transformer — новая архитектура нейросетей для работы с последовательностями / Семен Козлов, 2017 год - <https://habr.com/ru/users/sim0nsays/>
7. Автоматическая обработка текстов на естественном языке и компьютерная лингвистика: учеб. пособие / Большакова Е.И., Клышинский Э.С., Ландэ Д.В., Носков А.А., Пескова О.В., Ягунова Е.В. — М.: МИЭМ, 2011. — 272 с.
8. Компьютерная лингвистика / А. Н. Баранов // Большая российская энциклопедия : [в 35 т.] / гл. ред. Ю.С. Осипов. — М.: Большая российская энциклопедия, 2004—2017.
9. Боярский К.К. ВВЕДЕНИЕ В КОМПЬЮТЕРНУЮ ЛИНГВИСТИКУ/ К.К. Боярский// Учебное пособие. – СПб: НИУ ИТМО, 2013. – 72 с.
10. Generating music with AI (or transformers go brrrr) / Alexander Mamaev / 2021 год - <https://alxmamaev.medium.com/generating-music-with-ai-or-transformers-go-brrrr-3a3ac5a04126>

K.I. Abrosimov, A.S. Surkova

**MUSIC GENERATION ALGORITHM BASED ON ABC NOTATION AND
DISTRIBUTIVE SEMANTICS**

Nizhny Novgorod state technical university n.a. R. E. Alekseev

In this paper, we propose an algorithm for generating music, which is based on the representation of musical compositions in ABC notation, using the Word2Vec distributive semantics model, a modern neural network Architecture-Transformer. The results showed that the algorithm generates music no worse than the usual random mixing of prepared musical fragments, and meaningful melodic and rhythmic constructions are obtained.

Keywords: music generation, ABC notation, word2vec model, transformer neural network, finite state machines.

СРАВНЕНИЕ АЛГОРИТМОВ ТЕМАТИЧЕСКОГО МОДЕЛИРОВАНИЯ ДЛЯ ЛИТЕРАТУРНЫХ ТЕКСТОВ

Нижегородский государственный технический университет им. Р.Е. Алексеева

В данной работе исследованы и программно реализованы два метода тематического моделирования: латентно-семантический анализ LSA (Latent Semantic Analysis) и латентное размещение Дирихле LDA (Latent Dirichlet Allocation). В рамках статьи проведены эксперименты, позволяющие оценить скорость и качество каждого из рассмотренных алгоритмов. Для тестирования была выбрана коллекция документов, состоящая из произведений Н.В. Гоголя. Целью построения данной тематической модели являлось выявление основных тематик выбранных произведений.

Ключевые слова: компьютерная лингвистика, topic modeling, BTM, LDA, LSA.

Введение

Тематическое моделирование - способ построения модели коллекции текстовых документов, которая определяет, к каким темам относится каждый из документов. С помощью перехода из пространства терминов в пространство найденных тематик становится возможным решение таких задач, как тематический поиск, классификация и аннотация коллекций документов и новостных потоков. В основе алгоритмов тематического моделирования лежат статистические методы, анализирующие слова в текстах, с целью обнаружения тем, скрытых в этих текстах, а также взаимосвязи этих тем. [1].

При построении вероятностной тематической модели (BTM) предполагается, что коллекция документов - это последовательность терминов, выбранных случайно. Поскольку документ или термин может относиться одновременно ко многим темам с различными вероятностями, можно считать, что BTM осуществляет «мягкую» кластеризацию документов и терминов по кластерам-темам. Одним из основных приложений тематического моделирования является информационный поиск.

Наиболее популярные в настоящий момент методы тематического моделирования можно разделить на две основных группы - алгебраические и вероятностные. К алгебраическим моделям относятся стандартная векторная модель текста (VSM) и латентно-семантический анализ (LSA). Среди вероятностных моделей наиболее популярными являются вероятностный латентно-семантический анализ (pLSA) и латентное размещение Дирихле (LDA).

Постановка задачи тематического моделирования

Пусть D - коллекция текстовых документов, W - словарь всех встречающихся в них терминов (слов или словосочетаний). Каждый документ $d \in D$ представляет собой последовательность n_d терминов $(w_1, w_2, \dots, w_{n_d})$ из словаря W . Термин может повторяться в документе несколько раз.

Считаем, что существует конечное множество тем T , и каждое употребление термина w в каждом документе d связано с некоторой темой $t \in T$, которая не известна. Коллекция документов рассматривается как множество троек (d, w, t) , выбранных случайно. Документы $d \in D$ и термины $w \in W$ являются наблюдаемыми переменными, тема $t \in T$ является латентной (скрытой) переменной.

При построении тематической модели считаем, что порядок терминов в документах не важен для выявления тематики. Тематику документа можно узнать даже после перестановки терминов. Такое предположение называют гипотезой «мешка слов»

(bagofwords). Порядок документов в коллекции также не имеет значения - это предположение называют гипотезой «мешка документов».

Построить тематическую модель коллекции документов D - значит найти множество тем T , распределения $p(w|t)$ для всех тем $t \in T$ и распределения $p(t|d)$ для всех документов $d \in D$. Можно также говорить о задаче совместной «мягкой» кластеризации множества документов и множества слов по множеству кластеров-тем. Мягкая кластеризация означает, что каждый документ или термин не жёстко приписывается какой-то одной теме, а распределяется по нескольким темам. [2]

Предварительная обработка текстов

Для построения тематической модели, необходимо произвести предварительную обработку текстов, которая включает в себя несколько этапов. Сначала рекомендуется выполнить очистку текстов от знаков пунктуации, лишних пробельных символов и цифр. Следующим этапом произведем токенизацию- разбиение текста на текстовые единицы, например, слова или предложения.

Тексты на естественном языке имеют большое количество слов, которые не несут информации о данном тексте. Такие слова называют шумовыми. Например, к ним можно отнести предлоги, частицы, союзы. Для того, чтобы построить точную тематическую модель на этапе предварительной обработки текстов нужно удалить такие слова. На последнем этапе необходимо привести каждое слово к основе, одинаковой для всех его грамматических форм (произвести лемматизацию).

Все описанные выше модули обработки текстов были реализованы в программе, выполняющей предварительную обработку данных.

Латентно-семантический анализ

Латентно-семантический анализ (LSA, LatentSemanticAnalysis) - это метод для извлечения контекстно-зависимых значений слов при помощи статистической обработки больших наборов текстовых данных. LSA был запатентован в 1988 году.

LSA работает с векторным представлением типа «мешка слов» текстовых единиц. Текстовый корпус представляется в виде числовой матрицы - слово-текст, строки которой соответствуют словам, а столбцы - текстам. Объединение слов в темы и представление текстов в пространстве тем осуществляется путем применения к входной матрице одного из матричных разложений. Наиболее популярными являются: сингулярное разложение и факторизация неотрицательных матриц. [4]

Согласно теореме о сингулярном разложении, любая вещественная прямоугольная матрица может быть разложена на произведение трех матриц:

$$A = U * S * V^T$$

Здесь матрицы U и V^T - ортогональные, а S - диагональная матрица, значения на диагонали которой называются сингулярными значениями матрицы A . V^T -транспонированная матрица.

Если в матрице S оставить только k наибольших сингулярных значений, а в матрицах U и V^T - только соответствующие этим значениям столбцы, то произведение получившихся матриц $U_{lsa}, S_{lsa}, V^T_{lsa}$ будет наилучшим приближением исходной матрицы

A к матрице \tilde{A} ранга k :

$$A \approx \tilde{A} = U_{lsa} * S_{lsa} * V^T_{lsa}$$

Так как в качестве матрицы A использовалась матрица слово-текст, то матрица \tilde{A} , содержащая только k первых линейно независимых компонент A , отражает основную структуру различных зависимостей, присутствующих в исходной матрице. Структура зависимостей определяется весовыми функциями слов. Таким образом, каждое слово и текст представляются при помощи векторов в общем пространстве размерности k - пространстве гипотез. Сходство между любой комбинацией слов и/или текстов легко вычисляется при помощи скалярного произведения векторов. Как правило, выбор k зависит от поставленной задачи и подбирается эмпирически. [2]

Существенным недостатком метода LSA является значительное снижение скорости вычисления при увеличении объема входных данных.

Латентное размещение Дирихле

Латентное размещение Дирихле (LDA, LatentDirichletAllocation) - это порождающая модель, объясняющая результаты наблюдений с помощью неявных групп, что позволяет получить объяснение, почему некоторые части данных схожи [1]. Например, если наблюдениями являются слова, собранные в тексты, утверждается, что каждый текст представляет собой смесь небольшого количества тем и что появление каждого слова связано с одной из тем документа.

LDA впервые был представлен в качестве графической модели для обнаружения тем Дэвидом Блеем, Эндрю Нг и Майклом Джорданом в 2002 году. [6] С точки зрения модели LDA создание текста (предложений) происходит по следующей схеме:

1. Случайно выбрать для текста его распределение по темам θ_d
2. Для каждого слова в тексте:
 - (a) случайно выбрать тему из распределения θ_d , полученного на 1-м шаге.
 - (b) случайно выбрать слово из распределения слов в выбранной теме ϕ_t .

В рассматриваемом наборе текстов D каждый текст состоит из n_d слов.

Наблюдаемыми переменными являются слова в текста - w_d . Все остальные переменные - скрытые. Для каждого текста d переменная θ_d , представляет собой распределение тем в данном тексте. В классической модели LDA количество тем фиксировано и изначально задается параметром T .

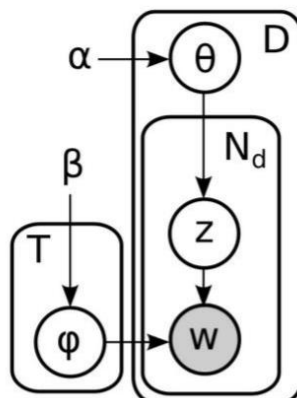


Рис. 1. Графическое представление модели LDA

Метод ЛРД основан на вероятностной модели и является развитием наивного Байесовского классификатора. Вероятность того, что в документе d встретится слово w описывается формулой:

$$p(d, w) = \sum_{t \in T} p(d)p(w|t)p(t|d)$$

При этом так же формулируются следующие предположения:

- векторы документов $\theta_d = (p(t|d): t \in T)$ порождаются одним и тем же

вероятностным распределением на нормированных $|T|$ -мерных векторах. Это распределение удобно взять из параметрического семейства распределений Дирихле $Dir(\theta, \alpha), \alpha \in R^{|T|}$.

- векторы тем $\phi_t = (p(w|t): w \in W)$ порождаются одним и тем же

вероятностным распределением на нормированных векторах размерности $|W|$. Это распределение удобно взять из параметрического семейства распределений Дирихле $Dir(\theta, \beta), \beta \in R^{|W|}$.

Здесь α и β – задаваемые вектора-параметры (гиперпараметры) распределения Дирихле.

LDA позволяет выявлять скрытую тематику документов на основе модели документа как «мешка слов». Предполагается, что существуют скрытые взаимосвязи между различными объектами. Семантическая близость различных объектов может оцениваться с помощью сравнения их тематических векторов. Основным недостатком распределения Дирихле является отсутствие убедительных лингвистических обоснований. С точки зрения текстов предположение о распределении Дирихле не является обоснованным. [7]

Экспериментальное сравнение работоспособности двух алгоритмов

В качестве исходного набора данных были выбраны литературные русскоязычные произведения Н.В. Гоголя («Невский проспект», «Нос», «Портрет», «Шинель», «Коляска»). Для программной реализации использовались библиотеки NLTK и rumorphu2 (необходимая для предварительной обработки текстов), а также Gensim (в этом пакете реализованы алгоритмы LDA, LSA). Целью построения тематической модели являлось выявление основных тем, содержащихся в выбранных произведениях, а также их визуализация с помощью «облаков» слов.

Построение модели напрямую зависит количества выбранных тем. Выбор этого параметра связан с понятием когерентности. Тема называется когерентной (согласованной), если термины, наиболее частые в данной теме, неслучайно совместно встречаются рядом в документах коллекции. Когерентность может оцениваться по сторонней коллекции (например, по Википедии), либо по той же коллекции, по которой строится модель. Для оценивания когерентности использовалась поточечная взаимная информация (pointwise mutual information, PMI). В результате анализа графика согласованности тем от их количества выявлено, что для выбранной коллекции документов оптимальное количество тем равно четырем.

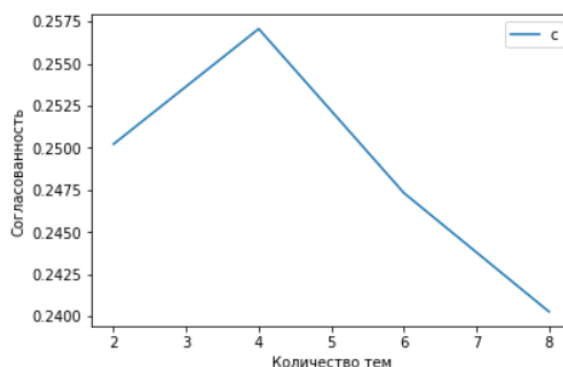


Рис. 2. График зависимости согласованности тем от их количества

Тематическая модель, построенная по алгоритму LSA, выглядит следующим образом:

```
[(0,
  '0.183*"человек" + 0.172*"акакии" + 0.168*"акакиевич" + 0.162*"лицо" + '
  '0.161*"глаз" + 0.150*"рука" + 0.146*"шинель" + 0.143*"нибыть" + '
  '0.142*"время" + 0.140*"нос"'),
(1,
  '-0.423*"акакии" + -0.415*"акакиевич" + -0.309*"шинель" + -0.243*"петрович" + '
  '+ 0.180*"художник" + 0.156*"портрет" + 0.153*"нос" + 0.116*"ковалев" + '
  '0.104*"глаз" + -0.099*"чиновник"'),
(2,
  '-0.487*"нос" + -0.364*"ковалев" + 0.220*"художник" + 0.206*"портрет" + '
  '+ -0.202*"иван" + -0.172*"маиор" + -0.151*"яковлевич" + -0.125*"ассессор" + '
  '0.122*"кисть" + -0.121*"коллежский"'),
(3,
  '-0.283*"шиллер" + -0.227*"пирог" + -0.205*"офицер" + -0.196*"невский" + '
  '+ -0.191*"проспект" + 0.180*"ковалев" + 0.180*"портрет" + 0.164*"нос" + '
  '0.140*"художник" + -0.138*"блондинка"')]
```

Рис. 3. Результат работы алгоритма LSA

В результате выполнения программы алгоритм LSA выделил 4 тематических кластера. Десять наиболее характерных слов для каждого класса можно увидеть выше. По этим термам достаточно просто определить тему каждого кластера. Выполнение алгоритма заняло 180ms. Однако при увеличении входных данных, время работы программы значительно возрастает.

Результат выполнения алгоритма LDA и визуализация в виде «облака» слов представлены на рис. 4 и 5. Каждое «облако» слов графически представляет ключевые слова определенной тематики. При отрисовке «облаков», учитывается вес каждого слова для выбранной темы: размер слова на графическом изображении пропорционален весу.

```
[(0,
  '0.009*"нос" + 0.008*"ковалев" + 0.006*"яковлевич" + 0.006*"иван" + '
  '0.004*"ассессор" + 0.004*"коллежский" + 0.002*"хлеб" + 0.002*"брить" + '
  '0.002*"извозчик" + 0.002*"статский"'),
(1,
  '0.006*"глаз" + 0.006*"художник" + 0.005*"портрет" + 0.004*"рука" + '
  '0.004*"лицо" + 0.004*"человек" + 0.004*"казаться" + 0.004*"нос" + '
  '0.003*"комната" + 0.003*"нибыть"'),
(2,
  '0.017*"акакии" + 0.017*"акакиевич" + 0.015*"шинель" + 0.008*"человек" + '
  '0.007*"время" + 0.007*"чиновник" + 0.007*"лицо" + 0.006*"петрович" + '
  '0.006*"значительный" + 0.005*"дело"'),
(3,
  '0.012*"офицер" + 0.009*"генерал" + 0.007*"чертокуцкии" + 0.006*"обед" + '
  '0.006*"рука" + 0.006*"господин" + 0.005*"пуф" + 0.005*"коляска" + '
  '0.005*"совершенно" + 0.005*"хорошии"')]
```

Рис. 4. Результат работы алгоритма LDA



Рис. 5. «Облака» ключевых слов для каждого кластера

Один проход данного алгоритма выполнялся 160 ms. Однако для улучшения качества построенной модели рекомендуется увеличивать количество итераций.

При сравнении результатов работы алгоритмов можно сделать следующий вывод: оба алгоритма тематического моделирования справились с выделением основных тем коллекции документов, однако LDA более полно описывает каждую тематику, выделяя характерные слова. Можно заметить, что ключевые слова, выделенные LSA, пересекаются в некоторых темах. Например, «Ковалев», «Акакий», «нос» встречаются сразу в трех темах. В модели LDA такого пересечения не наблюдается.

Также хочется отметить удобство специального инструмента визуализации результатов алгоритма LDA - pyLDAvis.

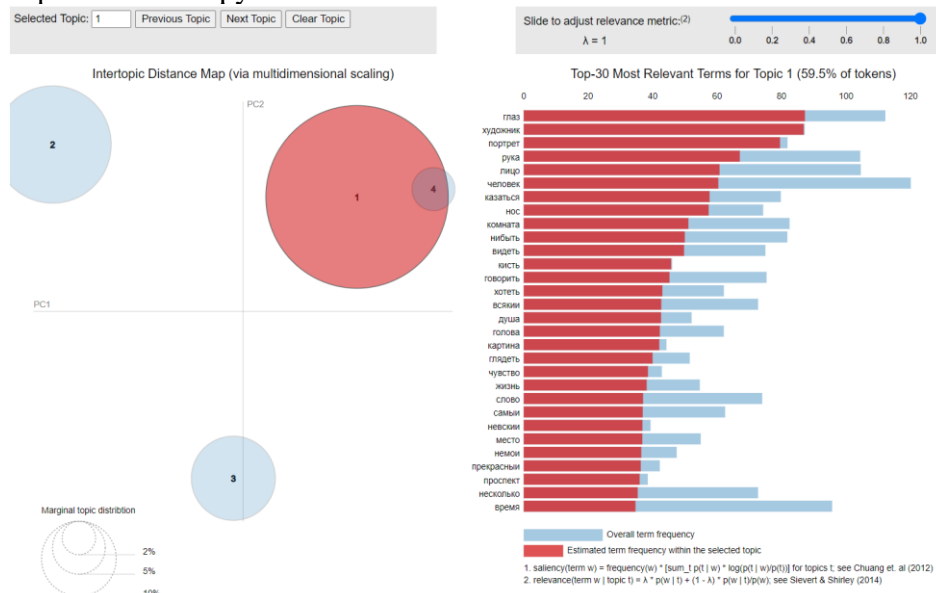


Рис. 6. Визуализация результатов алгоритма LDA с помощью pyLDAvis

Здесь каждый пузырь на левом графике представляет тему. Чем больше пузырь, тем больше распространена эта тема. Хорошая тематическая модель имеет довольно большие непересекающиеся пузыри, разбросанные по всей диаграмме. Если навести курсор на один из пузырей, справа обновятся полосы графиков. Эти слова являются

ключевыми словами, они формируют выбранную тему. Анализ этого графика показывает, что построенная модель является достаточно точной, однако существует перекрытие двух тем (1 и 4).

Заключение

В работе рассмотрена задача тематического моделирования текстовых данных. Для её решения изучены и реализованы два алгоритма: латентно-семантический анализ LSA и латентное размещение Дирихле LDA. Эксперименты показали, что на качество построенной модели значительное влияние оказывает заданное количество тем. Подбирая данный параметр следует учитывать влияние количества тем на их когерентность. В ходе экспериментов выяснилось, что оба алгоритма достаточно точно строят тематическую модель исходной коллекции документов, однако LDA работает с большей точностью. Также следует отметить значительное увеличение времени выполнения алгоритма LSA при увеличении входных данных. Для алгоритма LDA время выполнения возрастает менее резко при такой же нагрузке. Таким образом можно сделать вывод, что для литературных текстов целесообразнее использовать алгоритм латентного размещения Дирихле при тематическом моделировании.

Библиографический список

1. Билбро Ребекка, Бенгфорт Бенджамин. Прикладной анализ текстовых данных на Python. Машинное обучение и создание приложений обработки естественного языка. – 2019 г.
2. Воронцов К.В. Вероятностное тематическое моделирование // www.machinelearning.ru: web. - 2013.
3. Тематическое моделирование // Википедия, свободная энциклопедия. - URL: http://en.wikipedia.org/wiki/Тематическое_моделирование
4. Thomas Landauer, Peter W. Foltz, & Darrell Laham. Introduction to Latent Semantic Analysis // *Discourse Processes: journal*. — 1998. — Vol. 25. — P. 259—284.
5. Тематическое моделирование с помощью Gensim (Python) - URL: <https://webdevblog.ru/tematicheskoe-modelirovanie-s-pomoshhju-gensim-python/>
6. Визуализация тематического моделирования - URL: <https://webdevblog.ru/vizualizaciya-tematicheskogo-modelirovaniya-kak-predstavit-rezultaty-modelej-lda/>
7. Blei D.M., Ng A.Y., Jordan M.I. Latent Dirichlet Allocation // *Journal of Machine Learning Research*. - 2003. - 3(4–5). - Pp. 993-1022. - DOI: 10.1162/jmlr.2003.3.4-5.993.
8. Воронцов К., Потапенко А.А. Аддитивная регуляризация тематических моделей // *Доклады Академии наук*. - 2014. - Т. 456, № 3. - С. 268-271.

T. V. Lvutina, A. S. Surkova

COMPARISON OF THEMATIC MODELING ALGORITHMS FOR LITERARY TEXTS

Nizhny Novgorod state technical University. R. E. Alekseeva

In this paper, two methods of topic modeling are investigated and software implemented: Latent Semantic Analysis (LSA) and Latent Dirichlet Allocation (LDA). Within the framework of the article, experiments have been carried out to assess the speed and quality of each of the considered algorithms. For testing, a collection of documents was chosen, consisting of the works of N.V. Gogol. The purpose of building this thematic model was to identify the main topics of the selected works.

Keywords: computational linguistics, topic modeling, VTM, LDA, LSA.

МОДЕЛИ И МЕТОДЫ ДИАГНОСТИКИ СОСТОЯНИЯ СЛОЖНЫХ ДИНАМИЧЕСКИХ СИСТЕМ В УСЛОВИЯХ НЕПОЛНОТЫ ОБЪЕКТНЫХ ДАННЫХ

Нижегородский государственный технический университет им. Р.Е. Алексеева

Работа посвящена вопросу разработки методов диагностики состояния сложных динамических систем в условиях неполноты диагностической информации. Обоснована возможность использования математического аппарата дискретно-событийного моделирования. Предложена итерационная схема построения диагностических моделей и решения задачи диагностики, применительно к широкому классу задач.

Ключевые слова: Диагностика, дискретно-событийная модель, марковская модель, недетерминированный конечный автомат.

Введение

Сложные объекты, как живой, так неживой природы, представляют собой сверхсложные системы, компоненты которых находятся в непрерывном взаимодействии между собой и внешней средой. В данной работе под термином система будет подразумевать объект и его исследование будет выполняться в контексте процессов, протекающих в нем.

Процессы, протекающие в динамических системах, могут быть представлены в виде последовательности дискретных состояний S_i :

$$S = \bigcup S_i \mid i = \overline{0, N} \}, \quad (1)$$

Порядок следования состояний, в которых находилась исследуемая система, может быть существенно важен. Будем называть это упорядоченное множество состояний траекторией движения процесса, протекающего в системе.

Различие состояний системы возможно, как на основании количественных, так и на основании качественных оценок – множества параметров состояния X . То есть, множество источников данных о состоянии системы, определяющих X , может быть неоднородно по характеру оценки. Например, при анализе состояния металлоконструкций и сложных биологических систем оперируют и видеоинформацией, изображением объекта, и количественными оценками параметров его состояния. Определение состояния исследуемого объекта сводится к задаче многофакторного анализа. Для решения задачи мониторинга и диагностики состояния системы важно, что можно явно описать правила для определения факта перехода системы из одного состояния в другое, как некоторую функцию или решающее правило F :

$$F : X \rightarrow S, \quad (2)$$

где F – правило, позволяющее определить соответствие между набором значений параметров состояний системы X и соответствующим им состоянию S .

$$X = \{x_i \mid i = \overline{1, M}\}, \quad (3)$$

$$x_i = \bar{x}_i + \xi(t), \quad (4)$$

где X – множество параметров состояния исследуемого объекта, \bar{x}_i – усредненное значение i -го параметра состояния объекта, $\xi(t)$ – случайная составляющая.

Решающее правило или функция, определяющая зависимость между X и S , формирует решение, которое характеризуется определенной степенью точности и, соответственно, это отображение носит вероятностный характер. В динамических системах использование мгновенных значений параметров состояния системы может

приводить к значительным ошибкам в силу сложности определения характера изменения исследуемых параметров, который может иметь и стохастический характер. Поэтому используются, так называемые, суррогатные величины. В ряде случаев используется механизм усреднения значений параметров по времени. В этом случае, зависимость между X и S определяется следующим образом:

$$F : X \times X \rightarrow S . \quad (5)$$

Рассмотрим множество состояний S , описывающих процесс, протекающий в динамической системе. Состояние S_0 , в данной модели, соответствует начальному состоянию системы, а состояние S_N – конечному (финальному) состоянию. Величина N отражает количество состояний, через которое пройдет исследуемая система, пока не достигнет финального состояния.

Каждая компонента S_i множества S является отдельным множеством параметров состояния. Как правило, для сложных динамических систем, обладающих, в том числе, скрытыми параметрами состояния, количество параметров состояния не является определенной величиной. Этот факт отражается в нерегулярности мощности множества параметров состояния M_i для i -ого состояния системы:

$$S_i = \{x_{i,j} \mid i = \overline{1, M_i}\} , \quad (6)$$

где величина M_i характеризует мощность множества S_i , соответствующего i -му состоянию системы.

Стохастический характер величины размерности множества $\{x_{i,j}\}$ ставит перед исследователем задачу разработки алгоритмов классификации или распознавания состояния объекта. Эта задача связана с изучением статического образа системы. Для решения этой задачи принято использовать алгоритмы многофакторного анализа и решающие системы, основанные на искусственных нейронных сетях. Однако использование искусственных нейронных сетей связано с фазой обучения этих сетей, а для этого необходима достаточно большая по объему обучающая выборка. Как правило, объем таких выборок значительно превышает несколько тысяч, в то время как в случае решения медико-биологических задач, этот объем может не превышать нескольких сотен. Кроме того, сложность решения этой задачи обусловлена зашумленностью обрабатываемых данных, явным влиянием на результат систематических ошибок измерения параметров состояния исследуемой системы.

Следующим фактором, повышающим сложность задачи диагностики сложных систем, является процесс их взаимодействия с окружающей средой. В ряде случаев, это взаимодействие и его механизмы не только не изучены, но и сложно формализуемы. Что, в свою очередь, существенно усложняет аналитическое представление процесса функционирования системы.

Обозначим через W – множество параметров внешней среды, которые могут влиять на исследуемую систему, W^* – нечеткое подмножество параметров, оказывающих существенное влияние на исследуемую систему, а множество, описывающее процесс изменения значений параметров состояния через X^* : $X^* = X \times X$.

Тогда выражение (2) примет вид:

$$F : X^* \times W^* \rightarrow S \quad (7)$$

Модель (1) является простейшей дискретно событийной моделью. Она описывает только сами состояния, в которых находилась исследуемая система, но не содержит информации о контексте исследуемого процесса. Дальнейшим развитием данной модели можно считать модели конечных автоматов. Конечные автоматы позволяют оперировать такими понятиями, как вероятность перехода из одного состояния в другое, что дает возможность описывать отдельные участки траектории движения исследуемой системы.

Причинно-следственные связи для процессов и событий, протекающих в системе, отображаются в явном выделении отдельных подмножеств S_j^* :

$$S_j^* \subseteq S \times S \quad (8)$$

Фактически, величина S_j^* определяет траекторию состояний, которые испытывала j -ая система.

Диагностическая модель M – конечное множество детерминированных автоматных моделей, представляющих отображение $S \times X \rightarrow Y$:

$$M = \{A_i\}, \quad (9)$$

где A_i – конечный автомат, распознающий i -ое состояние, оперирующий данными не только i -го, но и предыдущих состояний. То есть, автомат A_i должен быть автоматом с памятью.

Конечным автоматом A называется пятерка $\langle S, X, Y, S_0, \delta, \lambda \rangle$, где S – конечное непустое множество состояний; X – конечное непустое множество входных сигналов (входной алфавит); Y – конечное непустое множество выходных сигналов (выходной алфавит); s_0 – начальное состояние; δ – функция переходов; λ – функция выходов.

$$A = \langle S, X, Y, S_0, \delta, \lambda \rangle, \quad (10)$$

$$\delta: S \times X \rightarrow S, \quad (11)$$

$$\lambda: S \times X \rightarrow Y \quad (12).$$

Выходные сигналы Y определяют качественные оценки состояния исследуемой системы. Они могут быть связаны со значением целевой функции системы. Следовательно, задача многофакторного анализа состояния такого класса систем может быть сведена к задаче синтеза конечного множества автоматов, позволяющих решать задачи распознавания отдельных состояний системы и автомата – арбитра, систематизирующего результаты, полученные от используемых автоматов.

Описание метода

При решении задачи диагностики могут решаться две задачи: определение исходного состояния системы на основании данных о процессе и определение конечного состояния. Он сводятся к определению функций перехода и функций выхода в выражении 10. Обе задачи оперируют данными о траектории движения системы и, как было сказано ранее, частично относятся к классу задач многофакторного анализа. Отличие от классической задачи многофакторного анализа состоит именно в необходимости анализа траектории движения системы, что возможно только при использовании моделей конечных автоматов с памятью. Рассмотрим траекторий движения двух систем A и B - S_A^* и S_B^* . Очевидно, что

$$S_A^* \subset S \times S, \quad S_B^* \subset S \times S. \quad (13)$$

При том, что

$$S_A^* \neq S_B^* \quad \text{OR} \quad S_A^* = S_B^*. \quad (14)$$

– финальные состояния систем A и B могут совпадать или различаться. Но результат, формируемый конечным автоматом с памятью, используемый для решения задачи диагностики процессов, протекающих в исследуемых системах, должен различаться, так как имеет место именно качественно различие процессов, протекающих в исследуемых системах.

Дискретно-событийная модель событий, возникающих в сложных динамических системах, является отражением причинно-следственных связей между процессами, протекающими как внутри системы, так и являющимися сложными процессами взаимодействия системы с внешней средой, в том числе информационной. Для систем, не обладающих свойством выделенной внутренней обратной связи, такие процессы, как правило, однозначно формализуются для некоторого события, процесса, являющегося причиной, указываются одно или несколько зависимых от него событий, следствий. Каждое из которых, в свою очередь, может послужить причиной возникновения других процессов. Такие процессы, как правило, описываются в виде марковских моделей, для которых можно явно вычислить вероятности переходов.

Рассмотрим задачу диагностики состояния систем, обладающих свойством самоорганизации и/или самовосстановления. Жизненный цикл работы такого класса сложных, динамических открытых систем упрощенно можно представить в виде размеченного графа состояний [2]:

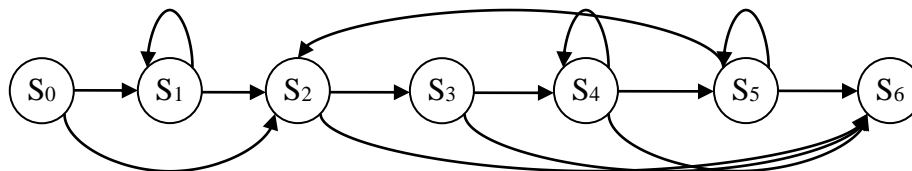


Рис. 1. Обобщенная графовая модель процессов, протекающих в системе

В графе, представленном на рис.1, S_0 – состояние системы на момент времени, соответствующее созданию новой системы, не содержащей внутренних дефектов; S_1 – состояние, соответствующее «приобретению» системой качественных изменений, отличающих ее от первоначального состояния (приобретение новых качеств); S_2 – состояние, соответствующее проявлению в работе системы некоторого дефекта, физического или иного деструктивного воздействия, который оказывает существенное влияние на процесс функционирования рассматриваемой системы; S_3 – состояние, соответствующее проявлению в работе системы «вторичных» повреждений, являющихся последствием «изменений свойств» системы, приобретенных в состояниях S_1 и/или S_2 ; S_4 – состояние системы, после частичного или полного восстановления работоспособности, обусловленное наличием в системе свойств и функций самовосстановления, адаптивности; S_5 – состояние системы, после процесса адаптации к изменению своих свойств; S_6 – конечное состояние, связанное с завершением функционирования системы, которое может быть обусловлено как физическим разрешением системы, так и завершением времени ее эксплуатации.

Подграф $\langle S_2, S_3, S_4 \rangle$ соответствует процессу качественного изменения характеристик системы. В результате данных изменений, система может либо частично сохранить, либо утратить возможность дальнейшего нормального функционирования. Поэтому состояние S_5 соответствует дальнейшему функционированию системы, претерпевшей некоторые качественные изменения и, как правило, такие изменения не являются положительными для ее дальнейшего функционирования. Дуги $\langle S_2, S_6 \rangle$, $\langle S_3, S_6 \rangle$, $\langle S_4, S_6 \rangle$ соответствуют досрочному прекращению функционирования системы.

Ориентированный граф, представленный на рис.1, соответствует дискретно-событийной модели процессов, протекающих в сложных динамических системах с дискретным набором состояний. В то же время, он может быть применен к описанию процессов в системах, описываемых при помощи дифференциальных уравнений [1, 3], если оперировать понятиями фазы развития процесса.

Применительно же к медицинским задачам можно сказать, что состояние S_2 соответствует событию проявления некоторого заболевания, которое может стать причиной возникновения ряда вторичных, зависимых от него, заболеваний – состояние S_3 . Вторичные поражения, как некоторые процессы, могут вызывать осложнение основного заболевания – состояние S_4 .

Задача диагностики, в этом случае, сводится к анализу подграфа $\langle S_2, S_3, S_4 \rangle$ и определению наиболее вероятного исхода процесса адаптации системы: состояний S_5 и/или S_6 . Причинно-следственные связи между событиями в системе могут быть формализованы, в том числе, средствами формальных логик, описывающими зависимость между качественными переходами системы из одного состояния в другое и отображены в

виде функции переходов δ . Это позволяет построить конечный автомат, соответствующий модели процесса, протекающего в системе.

В ряде случаев задача определения характера функции переходов δ сводится к вычислению корреляционной зависимости между парами параметров состояния исследуемой медико-биологической системы [5, 6]. Однако корреляционные модели позволяют выявить зависимости только в виде линейных функций и не позволяют определить функциональные зависимости другого вида. Математический аппарат конечных автоматов позволяет строить комбинированные решающие правила, определяющие функции переходов δ , оперируя, в том числе, и понятиями концептов.

Выделим в модели, представлено на рис.1, подграф, соответствующий этапу, на котором выполняется задача мониторинга и диагностики. Это подграф $\{S_2, S_3, S_4, S_5, S_6\}$.

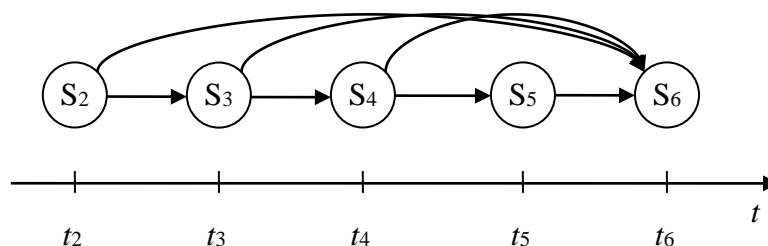


Рис. 2. Часть графовой модели процессов, для которой строится автоматная модель

В качестве целевой функции, исследуемой системы, определим сохранение системой возможности полной или частичной работоспособности. Это позволяет существенно сузить класс целевых состояний, используя всего три домена, соответствующие понятиям: полное сохранение работоспособности, частичная сохранение работоспособности и утрата работоспособности.

Заключение

Все вышеперечисленное однозначно позволяет использовать математический аппарат конечных автоматов для представления моделей процессов, протекающих внутри системы. Наличие скрытых и плохо формализуемых зависимостей между процессами, протекающими внутри системы и процессами, протекающими вне системы, но оказывающими на нее существенное влияние, обуславливает переход к модели не детерминированных конечных автоматов.

Дополнительный интерес при решении задачи диагностики представляет задача выделения концептов, описывающих траектории переходов исследуемых систем, то есть последовательности состояний, двигаясь по которым, система достигает определенных качественных финальных состояний, таких, как полное или частичное разрушение системы, сохранение или утрата функциональных качеств и другое. При таком подходе решение задачи диагностики может быть сведено к решению задачи распознавания образа процесса, использующего в качестве основных данных информацию о таких траекториях в контексте сопутствующей информации.

Библиографический список

1. Месарович М. Общая теория систем: математические основы / М. Месарович, Я. Такахага. – М.: Мир, 1978. – 312 с.
2. Мартынов Д. Разработка модели функционирования системы мозгового кровообращения при тяжелой черепно-мозговой травме / Д.С. Мартынов // Материалы VI международной научно-практической конференции "Информационные управляющие системы и технологии" (ИУСТ-ОДЕССА-2017) 20 - 22 сентября 2017 г., стр.271 – 273.
3. Неймарк Ю. Динамические модели теории управления / Ю.И. Неймарк, Н.Я. Коган, В.П. Савельев // М.: Наука. Главная редакция физико-математической литературы, 1985. – 400 с.
4. Ercole Ari. Visualisation of the 'Optimal Cerebral Perfusion' landscape in severe traumatic brain injury / Ari Ercole, Peter Smilevski, Robin Wesselink, Jan Elting, Joseph Donnelly, Marek Czosnyka, and Maurits Matasha // 16TH International Symposium on Intracranial Pressure & Neuromonitoring (IPC). In conjunction with the 6th Annual Meeting of the Cerebral Autoregulation Research Network. June 28 – July 2, 2016. Cambridge, USA. P.251-252.
5. Trofimov Alex. The Cerebrovascular Time Constant in Severe Polytraumatized Patients / Alex Trofimov, Dmitriy Martynov, Vladislav Pavlov, Vera Grigiryeva // Brain Edema 2014. The 16th International Conference on Brain Edema and Cellular Injury. The 3rd International Conference on Preconditioning for Neurological Disorders/ Huntington Beach, California September 27-30, 2014.
6. Trofimov Alex. The Cerebral Arterial Compliance in Severe Head Injury / A. Trofomov, A. Abashkin, D. Martynov, G. Kalenyev, M. Yuriev, V. Grigoryeva // 16TH International Symposium on Intracranial Pressure & Neuromonitoring (IPC). In conjunction with the 6th Annual Meeting of the Cerebral Autoregulation Research Network. June 28 – July 2, 2016. Cambridge, USA. P.147-149

D.S Martynov

MODELS AND METHODS FOR DIAGNOSTICS OF THE STATE OF COMPLEX DYNAMIC SYSTEMS UNDER CONDITIONS OF INCOMPLETE OBJECT DATA

Nizhny Novgorod state technical university n. a. R.E. Alekseev

The work is devoted to the development of methods for diagnosing the state of complex dynamic systems under conditions of incomplete diagnostic information. The possibility of using the mathematical apparatus of discrete-event modeling has been substantiated. An iterative scheme for constructing diagnostic models and solving the diagnostic problem is proposed for a wide class of problems.

Keywords: Diagnostics, discrete-event model, Markov model, non-deterministic finite state machine.

СУЩЕСТВУЮЩИЕ РЕШЕНИЯ РАСПОЗНАВАНИЯ ОБЪЕКТОВ НА ИЗОБРАЖЕНИИ

Нижегородский государственный технический университет им. Р. Е. Алексеева

В статье изложены одни из существующих алгоритмов, предназначенных для распознавания объектов, применяемые на практике.

Ключевые слова: распознавание объектов, вейвлет Хаара, гистограмма ориентированных градиентов, локальные бинарные шаблоны, компьютерное зрение.

Введение

В настоящее время существует множество алгоритмов для распознавания различных объектов, которые повседневно применяются в жизни многих людей и в различных сферах деятельности. Данная статья посвящена исследованию таких алгоритмов как: гистограммы направленных градиентов (Histogram of Oriented Gradients, HOG), локальные бинарные шаблоны (Local Binary Patterns, LBP), признаки Хаара (Haar wavelet).

Задача заключается в том, чтобы оценить эффективность алгоритмов, показать положительные и отрицательные стороны.

Примитивы Хаара

Каскад Хаара — способ обнаружения объектов на изображении, основанный на машинном обучении, идея которого была предложена в статье за авторством Пола Виолы (Paul Viola) и Майкла Джонса (Michael Jones) [1-2]. Принимая на вход изображение, обученный каскад Хаара определяет, есть ли на нем искомый объект, т.е. выполняет задачу классификации, разделяя входные данные на два класса (есть искомый объект, нет искомого объекта). Правильно обученный каскад Хаара имеет хорошую скорость выполнения классификации, а также неплохую устойчивость к отклонениям разного рода. [4-6]

Признак – отображение $f: X \Rightarrow D_f$, где D_f — множество допустимых значений признака. Вектор признаков $x = (f_1(x), \dots, f_n(x))$ называется признаковым описанием объекта $x \in X$, если заданы признаки f_1, \dots, f_n . Признаковые описания допустимо отождествлять с самими объектами. При этом множество $X = D_{f_1} * \dots * D_{f_n}$ называют признаковым пространством. Признаки делятся на следующие типы в зависимости от множества D_f :

- бинарный признак, $D_f = \{0,1\}$;
- номинальный признак: D_f — конечное множество;
- порядковый признак: D_f — конечное упорядоченное множество;
- количественный признак: D_f — множество действительных чисел.

Для решения прикладных задач с разнотипными признаками подходят далеко не все методы.

В стандартном методе Виолы – Джонса используются прямоугольные признаки, изображенные на рисунке ниже, они называются примитивами Хаара (рис. 1):

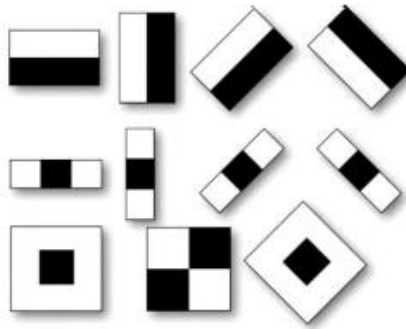


Рис. 1. Примитивы Хаара

В расширенном методе Виолы – Джонса, используемом в библиотеке *OpenCV* используются дополнительные признаки (рис. 2):

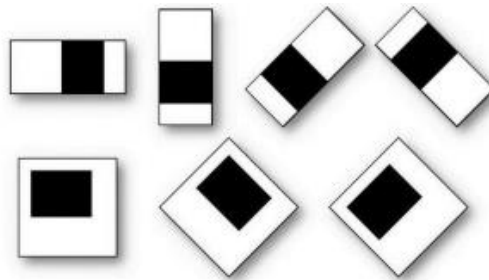


Рис. 2. Дополнительные примитивы Хаара

Вычисляемым значением такого признака будет $F = X - Y$, где X – сумма значений яркостей точек закрываемых светлой частью признака, а Y – сумма значений яркостей точек закрываемых темной частью признака.

Этот способ очень быстр в реализации, интуитивен и досконально известен. Алгоритм неустойчив при смене освещения, изменения масштаба или повороте изображения или часть изображения – изменяющийся фон, также очень маленькая скорость работы - если нужно обнаружить область $n * n$ на изображении $m * m$, то количество операций будет пропорционально $n^2 * (m - n)^2$.

Для того, чтобы модель работала как можно лучше возможны следующие решения: для нейтрализации освещения использовать нормировку или переход к бинаризации области, изменение масштаба и небольшие повороты нейтрализуются изменением разрешения и корреляции, скорость оптимизируется путем поиска с большим шагом или при маленьком разрешении.

Гистограмма ориентированных градиентов

В основе данного алгоритма лежит предположение, что вид распределения градиентов интенсивности изображения позволяет достаточно точно определить наличие и форму присутствующих на нём объектов.

При описании фрагмента изображения оно разбивается на несколько небольших участков, далее называемых *ячейками*. В ячейках вычисляются гистограммы h_i направленных градиентов внутри точек. Обычно они объединяются в одну гистограмму $h = f(h_1, \dots, h_k)$, после чего она нормализуется по яркости (L_2 и L_1 норма):

$$h_{L_2} = \frac{h}{\sqrt{|h|_2^2 + \varepsilon}}, h_{L_1} = \frac{h}{|h|_1 + \varepsilon}, h_{\sqrt{L_1}} = \sqrt{h_{L_1}}$$

где ε – некоторая малая константа.

Таким образом, данный описатель содержит пространственную информацию о фрагменте и инвариантен к освещению.

При вычислении градиентов производится свертка изображения с ядрами $[-1,0,1]$ и $[-1,0,1]^T$, в результате чего образуются две матрицы D_x и D_y производных вдоль осей x и y соответственно. Эти матрицы используются для вычисления углов и величин (модулей) градиентов в каждой точке изображения.

Пусть множество углов $(-\pi, \pi]$ разбивается на n равных интервалов вида $(-\frac{k-1}{n}\pi, \frac{k}{n}\pi)$, где $k = \{1, \dots, n\}$. Каждому интервалу ставится в соответствие бин гистограммы. Тогда гистограмма ячейки заполняется так, что величина градиента в каждой внутренней его точке добавляется к величине бина, соответствующего интервалу, содержащему угол данного градиента.

Экспериментально было доказано [3], что HOG значительно превосходит вейвлеты. Любая значительная степень сглаживания перед вычислением градиентов повреждает результаты HOG, подчеркивая, что большая часть доступной информации об изображении исходит от резких краев в тонких масштабах, и что размытие этого в надежде уменьшить чувствительность к пространственному положению является ошибкой, однако сильная локальная нормализация контраста необходима для достижения хороших результатов, а традиционные схемы окружения центра - не лучший выбор. Лучшие результаты могут быть достигнуты путем нормализации каждого элемента (ребра, ячейки) несколько раз по отношению к различным локальным опорам и обработки результатов как независимых сигналов.

Локальные бинарные шаблоны

Локальные бинарные шаблоны [11] представляют собой бинарные коды определенной разрядности, используемые для классификации в компьютерном зрении. Локальные бинарные шаблоны впервые были предложены в 1996 г. для анализа текстуры полутоновых изображений [7]. При этом дальнейшие исследования показали, что ЛБШ инвариантны к изменениям в условиях освещения и поворотам изображения [7, 8].

Локальный бинарный шаблон (ЛБШ) представляет собой описание окрестности пикселя изображения в двоичном представлении. Для вычисления базового ЛБШ в некоторой точке изображения используется восемь пикселей ее окрестности, а значение интенсивности центрального пикселя принимается в качестве порога. Пиксели со значением интенсивности, большим или равным пороговому значению принимают значения, равные единице, остальные принимают значения, равные нулю (рис. 3). Таким образом, результатом операции является восьмиразрядный бинарный код, который описывает окрестность этого пикселя [7].

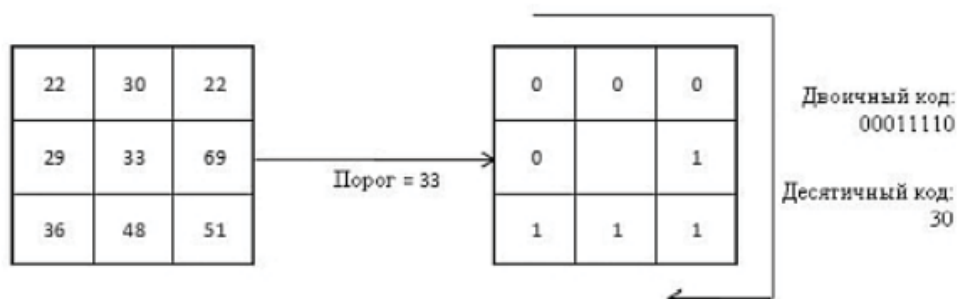


Рис. 3. Базовый ЛБШ

Для более гибкого анализа текстурных особенностей изображения используются круговая окрестность и билинейная интерполяция значений интенсивностей пикселей, которые позволяют построить расширенный ЛБШ (рис. 4) с произвольным числом точек P и радиусом R [7–10].

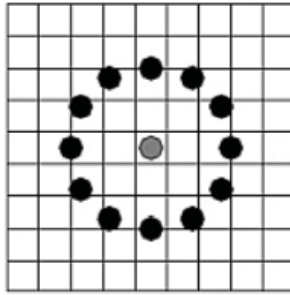


Рис. 4. Расширенный ЛБШ

Бинарные коды определенного вида несут в себе больше информации, чем остальные. ЛБШ называется равномерным (РЛБШ), содержащий не более трех серий “0” и “1” (например, 00011100, 011100000 и 11100001). Равномерные локальные бинарные шаблоны определяют только важные локальные особенности изображения: концы линий, грани, углы и пятна (рис. 5). Кроме того, они обеспечивают существенную экономию памяти (число РЛБШ равно $P * (P - 1) + 2$, число ЛБШ произвольного вида — 2^P).

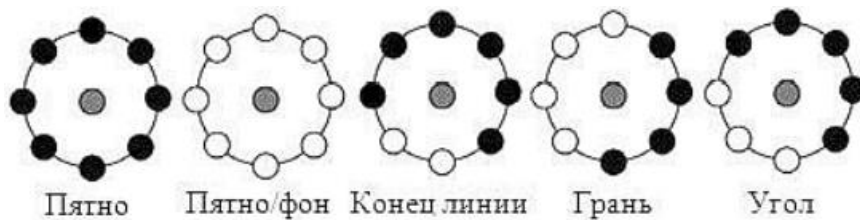


Рис. 5. Примеры локальных особенностей, детектируемых РЛБШ

Гистограмма ЛБШ.

Локальный бинарный шаблон вычисляется для каждого пикселя изображения, а затем строится гистограмма, в которой каждому равномерному коду ЛБШ соответствует отдельный столбец. Также формируется еще один дополнительный столбец, который содержит информацию обо всех неравномерных шаблонах. Изображения лиц могут быть представлены как набор всевозможных локальных особенностей, которые хорошо описываются с помощью ЛБШ. Однако гистограмма, построенная для всего изображения в целом, кодирует лишь наличие тех или иных локальных особенностей, но при этом не содержит никакой информации об их расположении на изображении. Для учета пространственного расположения изображение разбивается на подобласти, в каждой из которых вычисляется своя гистограмма ЛБШ (рис. 6).



Рис. 6. Разбитие изображения лица на подобласти

Путем конкатенации этих гистограмм строится общая гистограмма, учитывающая как локальные, так и глобальные особенности изображения. [8, 9]

Следует отметить, что качество работы системы существенным образом зависит от предобработки входных изображений, отсутствие этапа фильтрации приводит к снижению вероятностей правильной идентификации и верификации.

Библиографический список

1. Viola P. Rapid Object Detection using a Boosted Cascade of Simple Features / P. Viola, M.J. Jones // Proceedings IEEE Conf. on Computer Vision and Pattern Recognition (CVPR), 2001.
2. Viola P. Robust real-time face detection / P. Viola, M.J. Jones // Journal of Computer Vision. – 2004. – Vol. 57(2). – P. 137-154.
3. Dalal N., Triggs B. Histograms of Oriented Gradients for Human Detection // INRIA. — 2005.
4. Местецкий Л.М. Математические методы распознавания образов. – М.: МГУ, ВМиК, 2002-2004. – С. 42-44.
5. Sochman J. AdaBoost / J. Sochman, J. Matas // Center for Machine Perception, Czech Technical University, Prague, 2010.
6. Freund Y. A Short Introduction to Boosting / Y. Freund, R.E. Schapire. – Shannon Laboratory, USA, 1999. – P. 771-780.
7. Maenpaa T. The local binary pattern approach to texture analysis — Extensions and Applications. – Oulu University Press, 2003.
8. Shan C, Gong S., McOwan P. W. Facial expression based on local binary patterns: A comprehensive study // Image and Vision Computing. – 2009. (27).
9. Maturana D., Mery D., Soto A. Face recognition with local binary patterns, spatial pyramid histograms and naive bayes nearest neighbor classification // Proc. of the XXVIII International Conf. of the Chilean Computer Science Society, IEEE CS Society, 2009.
10. Ahonen T., Hadid A., Pietikainen M. Face recognition with local binary patterns. Lecture Notes in Computer Science, 2004.
11. Петрук В. И., Самородов А. В., Спиридов И. Н. Применение локальных бинарных шаблонов к решению задачи распознавания лиц. – Вестник МГТУ им. Баумана. Сер. «Приборостроение». 2011.

К. М. Ignakov

EXISTING SOLUTIONS FOR RECOGNIZING OBJECTS IN AN IMAGE

Nizhniy Novgorod state technical university n. a. R. E. Alekseev

Purpose: This article is devoted to the study of such algorithms as: Histogram of Directed Gradients (HOG), Local Binary Patterns (LBP), Haar features (Haar wavelet). The task is to evaluate the effectiveness of the algorithms, to show the positive and negative sides.

Keywords: object recognition, Haar wavelet, histogram of oriented gradients, local binary patterns, computer vision.

ПРОГРАММНАЯ СИСТЕМА ОТСЛЕЖИВАНИЯ КАЧЕСТВА ДОРОЖНОГО ПОКРЫТИЯ

Нижегородский государственный технический университет им. Р. Е. Алексеева

Целью работы является разработка приложения для проведения анализа качества дорожного покрытия, а также для отображения данных на устройстве пользователя

Ключевые слова: **дорожное** покрытие, приложение Android, акселерометр, сервер SQLite, анализ, качество

Введение

Одной из ключевых задач при планировании маршрута является надежность и безопасность. Разработка предназначена для анализа качества дорожного покрытия. В настоящее время существуют несколько способов получения анализа, но они имеют свои недостатки:

1. Дорогостоящее содержание и обслуживание инструментов для измерения
2. Большие временные затраты

Данные недостатки влекут за собой низкий уровень оценки качества дорожного полотна. Это приводит к увеличению скорости износа дорожного покрытия, что, в свою очередь, является одним из определяющих факторов дорожных аварий.

До сих пор одной из самых распространенных проблем России является качество дорожного покрытия. Оно влияет не только на ухудшение технических характеристик автотранспортного средства и его преждевременный выход из строя, но и на безопасность водителей и пассажиров. Согласно статистике, только за первые 6 месяцев 2020 года произошло 51.289 ДТП из-за нарушения обязательных требований к эксплуатационному состоянию автомобильных дорог.

Ежедневные поездки оказывают нагрузки на дороги, что приводит к износу дорожного полотна. Необходимо отслеживать состояние дорожного покрытия.

Данная разработка будет полезна для водителей, они смогут оценить по приложению состояние дорог и выбрать для них оптимальный и безопасный маршрут. Также у администрации города\района\области или региональных подразделений появится возможность отслеживания появления аварийных участков дороги и своевременного их устранения.

Результаты оценки контроля качества дорог нигде не публикуются для общего доступа, что делает использование данных в личных целях невозможным.

Программное обеспечение

Одним из самых важных модулей обнаружения и анализа вибраций отвечает акселерометр. Для этого мобильное устройство должно находиться в произвольном, но обязательно фиксированном положении в салоне автомобиля.

Устройство записывает маршрут своего передвижения, затем передает на сервер, после чего обрабатывает поступившие данные и готовит для отображения.

Можно выделить этапы действий над полученными данными. Сначала производится предварительная обработка данных, затем усреднение и объединение, в самом конце классификация.

Не менее важным является модуль местоположения устройства. Он предоставляет системе информацию о текущем местоположении устройства. Благодаря ему получаем данные о текущих координатах и скорости.

Также будет предусмотрена возможность:

1. Начало и остановка записи.
2. Выбор пользователя.
3. Синхронизация данных от нескольких пользователей.

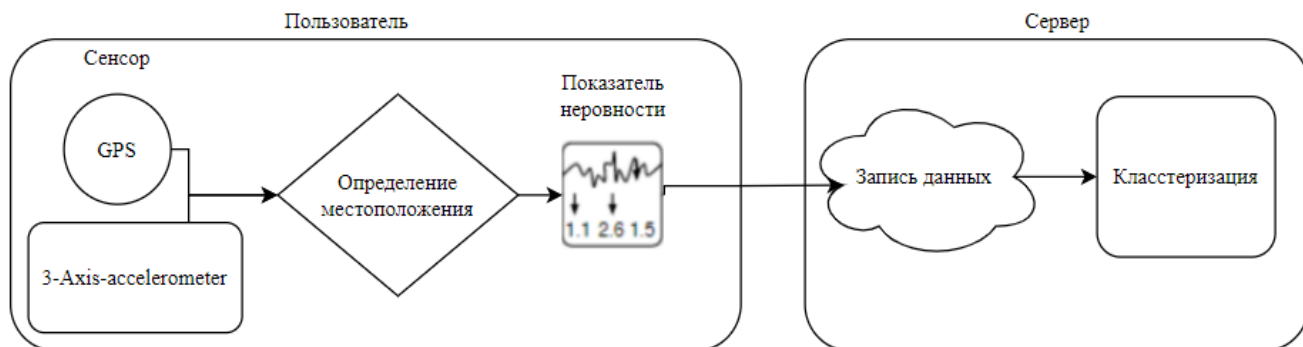


Рис. 1. Функциональная схема

Для принятия решение о категории качества дорожного полотна, установил категории качества и методы оценки принадлежности к ним. Необходимо вычислить характеристику каждой точки, характеристика – насколько сильно смартфон трясло в данной точке.

Была введена следующая шкала:

Таблица 1

Шкала состояния дороги

Состояние	Значение ускорения a , м/с
Отличное	$a \leq 1$
Хорошее	$1 < a \leq 2$
Удовлетворительное	$2 < a \leq 3$
Плохое	$3 < a \leq 4$
Ужасное	$a > 4$

Для вычисления a было принято решение использовать нахождение среднего значения среди всех полученных показателей за период измерения (за 0,5с):

$$a = \frac{1}{n} \sum_{i=1}^n x_i = \frac{1}{n} (x_1 + \dots + x_n),$$

Где a – усредненный показатель за период времени, n – число измерений за период времени, x_i – i -тое измерение уровня вибраций.

Характеристики каждого состояния:

1. Отличное – идеально ровная дорога, не имеющая видимых дефектов и неровностей.
2. Хорошее – целая ровная дорога, имеющая небольшие трещины, величина которых не превышает 10 мм.
3. Удовлетворительное – пригодная для езды, но имеющая заметные и ощущаемые дефекты, неровности, ямы глубиной до 20 мм.
4. Плохое – дорога, имеющая ямы глубиной более 20 мм, наблюдается частичное отсутствие твердого покрытия.
5. Ужасная – дорога опасная для передвижения на скорости 10-15 км\ч, большое количество и частота неровностей, ямы глубиной более 25 мм.

Каждой характеристике соответствует цвет точки:

1. Отличное – темно-зеленый.
2. Хорошее – светло-зеленый.

3. Удовлетворительное – желтый.
4. Плохое – оранжевый.
5. Ужасное – красный.

Точки изображаются на карте по координатам.

В качестве карты удобно использовать интернет-сервис «OpenStreetMap», т. к. он является некоммерческим проектом.

Точки изображаются на карте по координатам.

При загрузке отличающихся данных в одних координатах сохраняется последний вариант, т. к. данная информация является более актуальной.

Пример изображения состояния дорожного покрытия на карте показан на рис. 2.



Рис. 2. Состояние дорожного покрытия

Требования, предъявляемые к мобильному устройству пользователя. Для любого нового смартфона они являются несущественными. Можно выделить три основных требования:

1. Наличие подключения к Интернету.
2. Устройство должно иметь акселерометр.
3. Операционная система не ниже Android 4.4 KitKat.

Структурная схема

Автоматизированная система анализа качества дорожного покрытия состоит из двух частей (см. рис. 3):

- Серверная часть;
- Приложение для мобильного устройства.

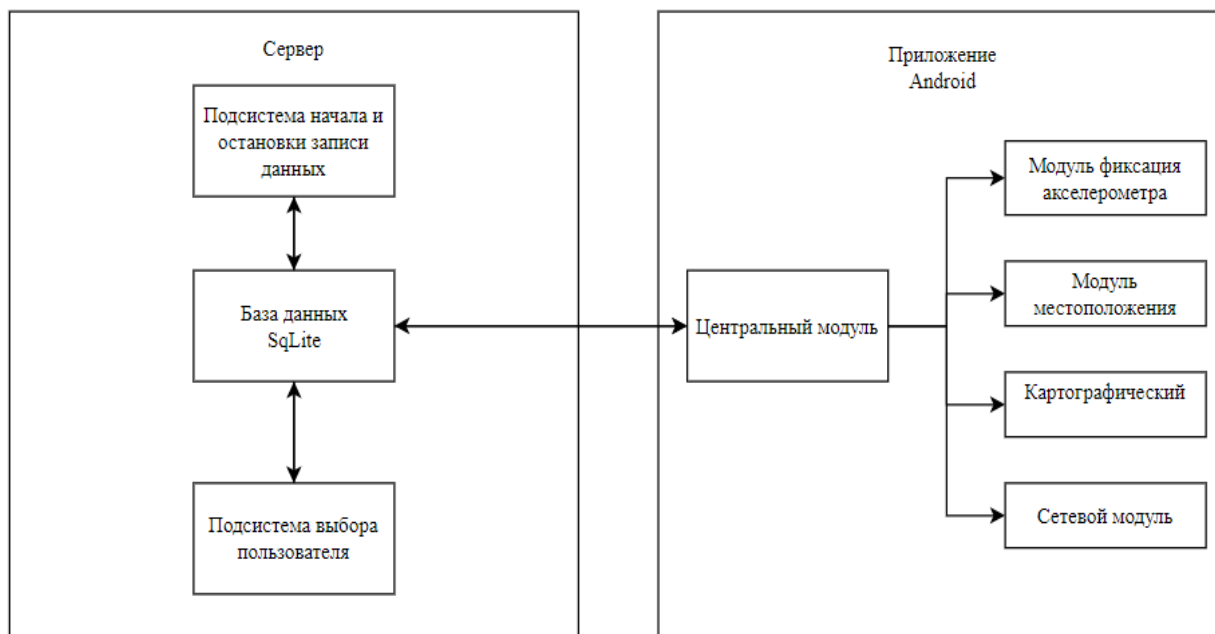


Рис. 3. Структурная схема

Заключение

Таким образом, был представлен прототип мониторинга качества дорожного полотна. Разработка состоит из мобильного приложения под Android и серверной программы SQLite.

Данная система имеет ряд преимуществ в сравнении с другими средствами анализа дорожного покрытия: удобство использования, высокая скорость анализа, общедоступность и компактность. В соответствии с этим разработка имеет большой потенциал для ее дальнейшего развития.

V.E. Gai, M.M. Bryzgalov

SOFTWARE SYSTEM FOR TRACKING THE QUALITY OF THE ROAD SURFACE

Nizhny Novgorod state technical university n.a. R.E. Alekseev

Purpose: the development is designed to analyze the quality of the road surface, as well as to display data on the user's device.

Design / methodology / approach: The Android application for the analysis of the road surface.

Findings: the author presented a prototype of monitoring the quality of the roadway.

Originality / value: this development will be useful for drivers, they will be able to assess the state of the roads using the application and choose the optimal and safe route for them, the results of the road quality control assessment are published for public access.

Keywords: road surface, Andoid app, accelerometer, SQLite server, analysis, quality.

АПАРТАНО-ПРОГРАММНАЯ СИСТЕМА УПРАВЛЕНИЯ ДВИЖЕНИЕМ МОБИЛЬНОГО РОБОТА НА ОСНОВЕ КИНЕКТ

Нижегородский государственный технический университет им. Р.Е. Алексеева

Целью работы является сборка робота из аппаратных средств: Kinect, Raspberry Pi, Arduino и программное написание алгоритма робота доставщика, задача которого заключается выйти из пункта А и достичь пункт Б, при этом учитывая имеются ли предметы или объекты, мешающие проезду робота. Если таковые имеются, то робот должен объехать препятствие, проверить достиг ли он точки Б, если нет, то движение в назначенном направлении продолжается.

Ключевые слова: робот, Kinect, компьютерное зрение, управление движением.

Введение

В последнее время стал активно развиваться ИИ (искусственный интеллект). Спектр задач, в которые привлекается ИИ достаточно разнообразен. В пример можно привести устройства для проведения операции пациенту, когда сам хирург находится в другом городе и нет возможности приехать. Можно привести пример связанный с сигнализацией, которая реагирует на дым и как только система его замечает, то дает сигнал пожара и старается его потушить. Есть такие проекты, в которых ИИ работает вместе с компьютерным зрением. Особенно это актуально в тот период времени, когда была объявлена пандемия. Люди и не подозревали о том, что им придется сидеть дома взаперти. Во время пандемии выходить людям было крайне не рекомендовано, а те, кто боялся заразиться сидели дома. И в данной ситуации было бы неплохо иметь какого-то курьера, работающего через компьютерное зрение, задача которого доставлять из пункта А в пункт Б. Если же встречается препятствие, то робот должен понимать, что ему что-то мешает проехать. Его задачей было бы объехать препятствие и продолжить свой маршрут. Данный робот курьер, помогал бы людям, как минимум закупать вещи, продукты и доставлял их, тем самым оставляя клиентов в безопасности.

Программные средства

libfreenect – открытое программное обеспечение, которое обеспечивает работу Kinect с Linux, Windows, Mac. Имеется вторая версия для обеспечения поддержки Kinect второго поколения. Имеется возможность получения RGB – изображения, запись аудио, управление мотором. Для работы с данным ПО необходимо установить libusb, cmake. Имеется подробная документация, работать можно на языках программирования: C, C++, Python, Java и другие. (рис. 1).



Рис. 1. OpenKinect

PyCharm – интегрированная среда разработки для языка программирования Python. Предоставляет средства для анализа кода, графический отладчик, инструмент для запуска юнит-тестов и поддерживает веб-разработку. Кроссплатформенная среда разработки, которая совместима с Linux, Windows, Mac. PyCharm Community Edition. (рис. 2).



Рис. 2. PyCharm

Аппаратные средства

Kinect (рис. 3) – это камера, выпущенная компанией Microsoft для консоли Xbox 360. Но нашлась группа энтузиастов, которые смогли взломать довольно легко данную камеру. Прошло некоторое время и для Kinect был выпущен официальный набор инструментов для разработчиков от Microsoft. Но и появились проекты OpenSource, которые дают легкость к подключению работая на платформах Linux, Windows, Mac. Технические характеристики: RGB-камера с помощью неё происходит цветное потоковое видео с частотой 30 кадров в секунду, и разрешением VGA, разрешение которого – 640x480. Инфракрасный проектор, он создает сетку из точек в пространстве. Камера, снимающая в инфракрасном режиме, которая регистрирует изображение сетки.



Рис. 3. Kinect

Rasbbery Pi (рис. 4)– компьютер небольшой размером. Rasbbery Pi выпускается в двух версиях – Model A и Model B. У Model A нет порта Ethernet, один порт USB 2.0 и 256 Мбайт RAM. Model B имеет порт Ethernet 10/100 Мбит/с, с двумя портами USB 2.0 и объем памяти в два раза больше, чем у модели Model A. Соответственно энергопотребление у двух моделей разное: Model A – 500 мА (2,5 Вт), Model B – 700 мА (3,5 Вт).



Рис. 4. Rasbbery PI

Arduino (рис. 5)- небольшая плата с собственным процессором и памятью. Длина и ширина платы – 69 мм x 53 мм. Расстояние между выводами соответствует стандарту 2,54 мм. Сама плата имеет 6 аналоговых входов и 14 цифровых входов общего назначения, кварцевый генератор на 16 МГц, также имеется два разъема, а именно силовой и USB, разъем ISPC необходим для внутрисхемного программирования и кнопка горячей перезагрузки. Чтобы плата стабильно работала её нужно подключить через встроенный USB разъем.



Рис. 5. Arduino

Алгоритм

У робота есть задача выйти из пункта А и добраться до пункта Б, при этом ему необходимо понимать, есть ли перед ним препятствие и если оно есть, то ему необходимо принять решение объехать его. С помощью устройства Kinect он будет распознавать, находится ли перед ним предмет, который мешает ему пройти. Если такой предмет или объект имеется, то робот принимает решение объехать его. Пользователь указывает откуда начинать путь, то есть пункт А и где путь заканчивается, то есть точка Б. Через Kinect робот смотрит, если препятствие имеется, ему необходимо его объехать, после чего происходит проверка на расстоянии от пункта Б робот остановился и изменил свое движение, чтобы объехать препятствие. Если робот объезжая достиг пункт Б, то продолжение движения не произойдет, если же робот так и не достиг пункта Б, он продолжит движение, до тех пор, пока не встретит новое препятствие, тем самым объезжая его и делает проверку, преодолел ли он препятствие и достиг ли конечного пункта.



Рис. 6. Робот



Рис. 7. Блок-схема алгоритма

Несомненно, данная разработка будет актуальна. Мир не стоит на месте и развивается, а технологии должны помогать и упрощать жизнь людей. Были представлены основные моменты и особенности работы робота, были определены программные

средства такие как libfreenect для получения доступа к Kinect, были отмечены моменты для корректной работы данного программного обеспечения, PyCharm – для комфортной работы программирования робота на языке программирования Python. Kinect, через который робот определяет если предмет и объект препятствующий его движению были рассмотрены его технические характеристики. Raspberry был описан, что из себя представляет, были рассмотрены особенности каждой модели и обозначены различия. Arduino рассмотрели каких размеров, что в нем имеется для работы. Также был рассмотрен алгоритм, приведена блок-схема к ней и подробное описание. Все выше перечисленное вполне достаточно для создания и реализации данного робота.

Библиографический список

1. Устройство Kinect URL:<https://xakep.ru/2011/09/21/56900/>
2. Что такое Arduino URL:<https://amperka.ru/page/what-is-arduino>
3. Что такое Raspberry Pi URL:https://ru.wikipedia.org/wiki/Raspberry_Pi
4. Открытый проект libfreenect URL:<https://github.com/OpenKinect/libfreenect>
6. Марк Лутцк. Программирование на Python.

D.V. Savin, V.E. Gai

KINECT-BASED APPARATUS-SOFTWARE SYSTEM FOR CONTROLLING THE MOTION OF A MOBILE ROBOT

Nizhny Novgorod state technical university n.a. R.E. Alekseev

Purpose: assembling and writing a program for a delivery robot working via Kinect.

Design/methodology/approach: assembly, programming, computer vision.

Findings: the author proposed hardware, software and algorithm.

Research limitations/implications: the robot can be used as a delivery man, a carrier.

Originality/value: the robot can work through the Kinect device, identify objects or objects that interfere with reaching the final destination, overcome them as a detour and if there are no obstacles can continue to move

Keywords: robot, Kinect, computer vision, motion control.

ПРОГРАММНАЯ СИСТЕМА РАСПОЗНАВАНИЯ ДОРОЖНЫХ ЗНАКОВ И РАЗМЕТКИ В СИМУЛЯТОРЕ WEBOTS

Нижегородский государственный технический университет им. Р.Е. Алексеева

Целью работы является создание системы распознавания дорожных знаков и разметки в симуляторе Webots.

Ключевые слова: система распознавания, дорожные знаки, разметка, контроль полосы, адаптивный круиз, симулятор Webots.

Введение

В современном мире одним из основных векторов развития науки и техники является автоматизация различных процессов во всех сферах человеческой деятельности и передача рутинных задач машинам и технике. Одна из таких задач - быстрое и безопасное передвижение.

Эта система достаточно необходима на сегодняшний день, т.к. с каждым годом растёт количество машин в городах. Из источника ГОУБДД МВД РОССИИ удалось выяснить, что из основных причин аварий почти 58% это несоответствие скоростному режиму. В связи с этим инженеры, стремясь повысить безопасность машин, разработали данную систему. Основная задача этой системы сводится к тому, чтобы отслеживать дорожные знаки и дорожное полотно и считывать с них информацию.



Рис. 1. Основные причины ДТП из источника ГОУБДД МВД РОССИИ
Распознавание дорожных знаков

Система распознавания дорожных знаков должна определять следующие знаки [1]:

- Ограничение скорости / Снятие ограничения скорости.
- Ограничение обгона / Снятие ограничения обгона.
- Участки с запретом на остановку.
- Начало и конец населённого пункта.
- Знак «Уступи дорогу».
- Запрет движения.

Конструкция и логика работы:

В системе распознавания дорожных знаков обязательно присутствуют:

1. Камера – с помощью неё бортовая электроника считывает дорожные знаки;

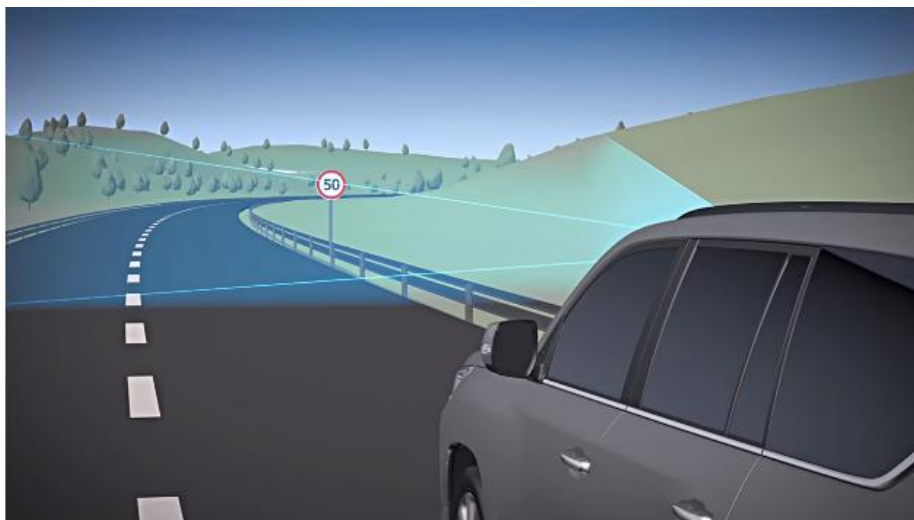


Рис. 2. Наглядный пример считывания дорожного знака [2]

2. Электронный блок управления - обрабатывает информацию, которая поступает с камеры, и принимает соответствующие решения, ориентируясь на собственную базу данных;

3. Дисплей – для того, чтобы на него выводилась информация о действующих ограничениях на конкретном участке пути.



Рис. 3. Информация на дисплее после распознавания знака (ограничение скорости 50 км/ч) [3]

Пошаговое описание логики обнаружения объектов [3]:

1. Камера анализирует окружающую среду и считывает данные о дорожных знаках.
2. Система обнаруживает форму, похожую на знак.
3. Распознавание цвета и наличия дополнительных символов.
4. Поиск соответствий в базе данных.
5. Информирование водителя через дисплей.

Последовательность распознавания типа знака [3]:

1. Определение формы: круг, прямоугольник, квадрат.
2. Анализ цветовой гаммы.
3. Считывание символов или надписей на знаке.
4. Сравнение фактической скорости с допустимой.

Использование таких систем распознавания знаков характеризуется рядом следующих преимуществ:

- Слежение за допустимой скоростью движения.
- Увеличение безопасности автомобиля, снижение вероятности аварии.
- Повышение комфорта использования транспортного средства.

Из недостатков данной системы можно выделить проблемы распознавания знаков в слепых зонах камеры, например, не при движении на дороге свыше двух полос, когда знаки расположены только на левой части дороги.

Распознавание разметки

Во время движения автомобиль должен находиться в своей полосе и не должен из нее выезжать, но на деле же водитель может отвлекаться от управления автомобилем, и машина может начать перемещаться на соседнюю полосу, что может привести к особенно печальным последствиям.

Для предотвращения этого люди придумали отслеживать позицию автомобиля в полосе и в случае, если он начинает уходить с нее, подавать сигнал водителю или применять более активные действия: поворачивать руль или применять тормоза на колесах противоположной стороны, чтобы не дать автомобилю покинуть свою полосу.

Принцип работы системы распознавания разметки:

Первый шаг работы алгоритма заключается в том, что картинка, передающаяся с камеры, переносит предобработку – фильтрация от шумов и векторизация изображение.



Рис. 4. На снимках показан результат комбинированной маской и затемнённым изображением [5]

После чего происходит обновление состояния линий дорожной разметки по данным из предыдущего шага.

И рисование обновленных линий и других объектов на исходном изображении.



Рис. 5. Обновленные линии на исходном изображении [5].

После распознавания дорожной разметки анализируется направление движение, если оно отклоняется от полосы движения производится сигнал (например, вибрация руля), если реакции от водителя не поступает, производится торможение для удержания автомобиля в полосе [6].

Заключение

Система распознавания дорожных знаков и разметки очень актуально в наше время, когда машин становится всё больше, таких «помощников» создают для большей безопасности на дорогах.

Были рассмотрены принципы работы системы распознавания дорожных знаков и разметки. И указал основные моменты на что стоит обратить внимание при изучении темы.

Библиографический список

1. Система распознавания дорожных знаков [Электронный ресурс]: naavtotrasse. - Режим доступа: <https://naavtotrasse.ru/encziklopediya/sistema-raspoznaniya-dorozhnyh-znakov.html>
2. Система распознавания дорожных знаков [Электронный ресурс]: YouTube.- Режим доступа: <https://www.youtube.com/watch?v=fU--WK1jEUK>
3. Система распознавания дорожных знаков: описание и принципы работы [Электронный ресурс]: ТехАвтоПорт. - Режим доступа: <https://techautoport.ru/sistemy-bezopasnosti/aktivnaya/sistema-raspoznaniya-dorozhnyh-znakov.html>
4. Принцип работы детектора [Электронный ресурс]: ЯндексДзен. - Режим доступа: <https://zen.yandex.ru/media/id/5aba894879885e29b6af5efe/vvedenie-v-opencv-primenitelno-k-raspoznaniyu-linii-dorojnoi-razmetki--5ad81aac79885e87ed4603a2>
5. Простой алгоритм распознавания дорожной разметки [Электронный ресурс]: REG.RU / Блог. - Режим доступа: <https://www.reg.ru/blog/simple-algorithm-for-road-marking-detection/>
6. Активная система удержания полосы движения [Электронный ресурс]: YouTube. - Режим доступа: <https://www.youtube.com/watch?v=RPuFwa66IpQ>.

V.E. Gai, S.S. Pegov

SOFTWARE SYSTEM FOR RECOGNIZING ROAD SIGNS AND MARKINGS IN THE WEBOTS SIMULATOR

Nizhny Novgorod state technical university n.a. R.E. Alekseev

Purpose: to create a system for recognizing road signs and markings in the Webots simulator.

Design/methodology/approach: a system for recognizing road signs in an image using machine learning methods.

Findings: the recognition system will be used in the Webots simulator.

Research limitations/implications: the system can be adapted for any high-definition camera, and installed in a car.

Originality/value: the system can be used for road sign recognition, marking recognition. In the future, it can be further developed and presented as an unmanned transport system.

Keywords: recognition of road signs, markings, traffic lanes, car, adaptive cruise control.

ПРОГРАММНАЯ СИСТЕМА УПРАВЛЕНИЯ АВТОМОБИЛЕМ В СИМУЛЯТОРЕ WEBOTS

Нижегородский государственный технический университет им. Р.Е. Алексеева

Целью работы является разработка программной системы управления автомобилем в симуляторе Webots, которая реализовывает безопасное движения автомобиля по многополосной дороге с множеством статических и динамических препятствий.

Ключевые слова: автопилотирование, программная система управления(CU), PID-регулятор, датчики расстояния, Webots

Введение

В настоящее время технологии автопилотирования транспортных средств являются одними из наиболее интересных, перспективных и массовых технологий, получивших развитие. По мнению некоторых источников, к 2035 году беспилотные автомобили могут составить до 25% от всех продаваемых в мире. Это поможет существенно снизить аварии на дорогах и на уменьшить число выпускаемых машинами выхлопных газов [1].

Зарождение идеи создания автомобилей с автопилотированием берет свое начало с момента, когда автомобильные конструкторы поняли, что электроника способна более эффективно справляться с блокировкой колес, в отличие от человека. После подтверждения данного умозаключения начали создаваться антиблокировочные системы тормозов (Anti-lock braking system) [2], которые получили большую популярность. Прогресс не стоял на месте, поэтому появились антипробуксовочные системы (Traction Control System) [2], которые контролируют оптимальное распределение момента между колесами, чтобы автомобиль не пробуксовывал. Далее появилась система контроля устойчивости автомобиля (Electronic Stability Program) [2], которая способна бороться со сносами и заносами, выборочно подтормаживая колеса и регулируя тягу двигателя. Позже автомобильные конструкторы пришли к выводу, что рулевое управление также можно доверить электронике, т.к. системы управления могли повторить один и тот же маневр с одинаковой точностью, а также провести анализ ситуации быстрее чем человеческий мозг.

Первые прототипы беспилотных автомобилей появились в 1980-х.: 1984 году проект Navlab [3], который включал серию автономных и полуавтономных автомобилей, справляющихся с ездой на бездорожье и автоматизированных автомагистралях, а также предотвращением столкновений на бездорожье и помощью водителю при маневрировании в многолюдных городских условиях.

К 2021 году существуют беспилотники, которые способны успешно передвигаться в городской многолюдной среде, а также на бездорожье без вмешательства человека. Лидерство в создании беспилотных автомобилей принадлежит таким компаниям, как TeslaMotors, Waymo и OttoGroup [3].

Важно сказать, что изучение технологий автопилотирования среди студентов имеет значительную роль, в дальнейшем развитии этих технологий. Это дает возможность увидеть преимущества и недостатки определенных методик и алгоритмов, а также показывает дальнейший потенциал в их использовании.

Программные средства

Для реализации системы управления (СУ) автомобиля необходимо определить список программный средств и инструментов для достижения цели.

Реализация СУ будет оформлена как проект в многоплатформенной десктопной программе Webots, которая служит для моделирования роботов. Она предоставляет полную среду разработки для программирования и симулирования роботов. Проект Webots [4] будет включать виртуальное пространство, в котором реализована сцена для решения задачи, а также модель машины, для которой будет написан контроллер, осуществляющий движение модели машины в зависимости от ситуации.

В качестве модели машины был взят среднеразмерный престижный седан компании Lincoln, Lincoln MKZ.



Рис. 1. Модель Lincoln MKZ, используемая для реализации СУ

В качестве среды проведения эксперимента над движением модели было выбрано виртуальное пространство HighwayDriving. В нем существует многополосная дорога с препятствиями и поворотами.

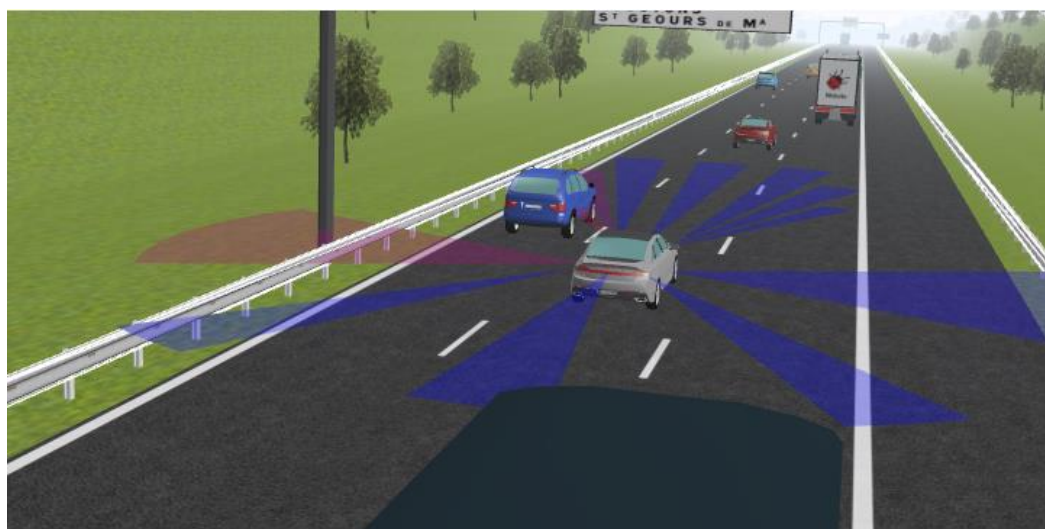


Рис. 2. Виртуальное пространство HighwayDriving, используемое для реализации СУ

Описание требований к алгоритму

Существует многополосная дорога, по которой движутся транспортные средства. На дороге имеются препятствия как статические (неровности дороги), так и динамические (другие транспортные средства). Машина с алгоритмом автопилотирования должна двигаться по крайней правой полосе вдоль отбойника за впереди едущим автомобилем, уметь обгонять другие автомобили, применяя ускорение, а также вписываться в повороты, применяя торможение.

Реализация

Для того чтобы реализовать алгоритм системы управления автомобилем, необходимо спроектировать основной контроллер, это программа, которая заставляет автомобиль двигаться в зависимости от ситуации.

Главными компонентами для реализации алгоритма являются датчики измерения расстояния [5], которые будут использоваться для предупреждения дальнейшего маневрирования. По рис. 3 можно проследить, что у модели автомобиля имеется 12 датчиков расстояния, но в работе не будут использованы все датчики.

Используемые датчики:

1. Right датчик – используется для определения расстояния до отбойника, чтобы знать, на какой полосе находится автомобиль, а также для определения наличия препятствия на полосах правее той, на которой находится автомобиль.

2. Front датчик – используется для определения расстояния до находящегося впереди препятствия. При анализе категории какого-либо препятствия (статическое или динамическое) будет приниматься решения либо об уменьшении скорости, либо о перестроении на соседнюю полосу.

3. Left датчик – используется для определения наличия препятствия на полосах, левее той, на которой находится автомобиль.

Датчики категории FrontLeft/Right – используются для определения наличия препятствий на соответствующих полосах для принятия решения о перестроении.

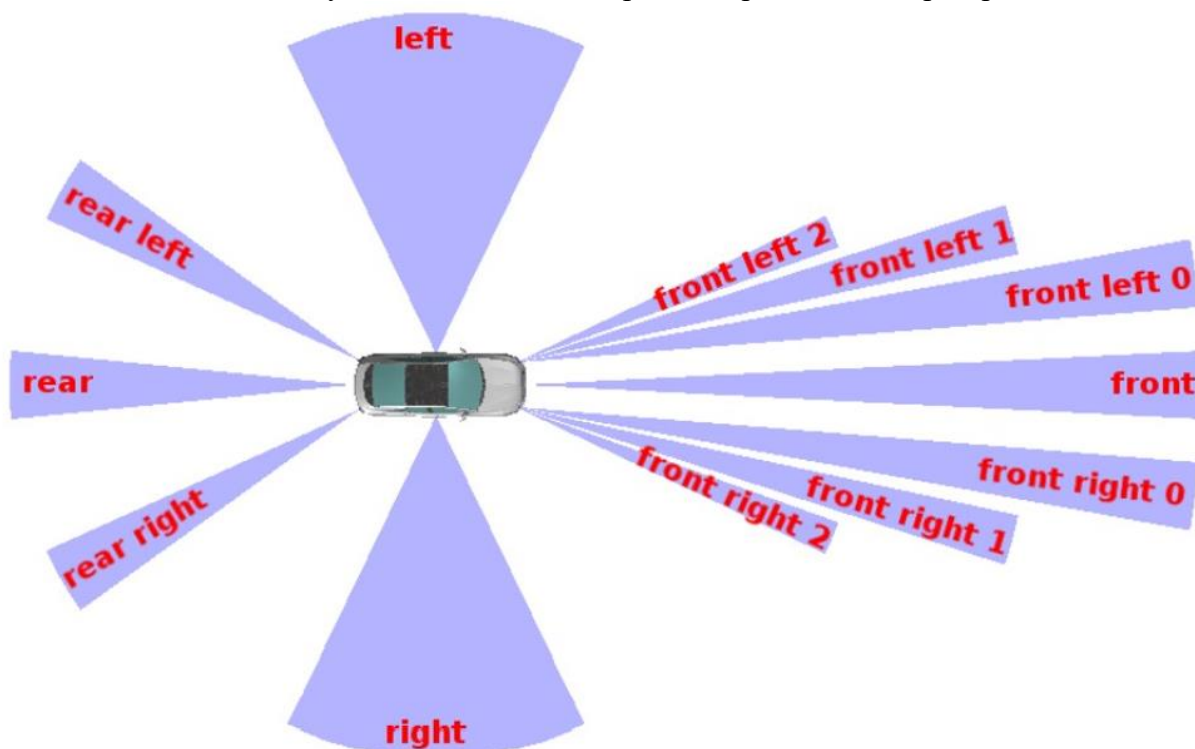


Рис. 3. Датчики расстояния для модели Lincoln MKZ

PID-регулятор [6, 7], который будет контролировать маневры. ВPID-регуляторе три составляющие, каждая из них будет определяться для конкретной цели: Пропорциональная составляющая [8] будет базовым углом поворота, интегральная составляющая [8] будет минимизировать ошибки, что не позволит увеличиваться колебаниям, дифференциальная составляющая [8] будет ускорять или замедлять повороты, что позволит производить движения в поворотах при извилисты дорогах наиболее точно.

Для вычисления каждой составляющей с правого датчика снимаются показания и вычисляется отклонение от желаемой линии (6 метров от отбойника соответствуют середине крайней правой полосы).

Все три составляющие и их коэффициенты, используемые для настройки точности выполнения маневров, должны быть отрегулированы относительно друг друга, а также переводиться в значения, которые может принять поворачивающее устройство (угол поворота управляющих колес). Меняя данные коэффициенты, можно получить необходимую конфигурацию соотношений, составляющих и перевести их результат в угол поворота. К тому же, данные коэффициенты позволяют настроить машину под конкретную трассу.

С помощью модуля Driver из библиотеки Vehicle, которая представляет собой средства для управления автомобилями, будут подаваться команды нажатия на педали газа и тормоза, выставляться круизная скорость и поворачиваться руль.

В алгоритме присутствует блок регулировки скорости. Датчик front будет определять расстояние до впереди находящегося препятствия. Показания, которые будут сняты с этого датчика переводятся в доли (делятся на максимальное расстояние, которое фиксирует данный датчик – 20), а полученный коэффициент умножается на определенную максимальную скорость. Поэтому если машина приближается к препятствию, то скорость начнет снижаться. Замедляться сам автомобиль не будет, т.к. скорость, которая была вычислена – круизная скорость, а замедление производится нажатием на педаль тормоза с определенной силой. Если реальная скорость выше круизной, тосила нажатия на тормоз будет определяться разницей скоростей, в эквиваленте 1 к 100% (полное нажатие на педаль тормоза). Если реальная скорость сравнялась с круизной, то нажатие на педаль тормоза прекращается.

Заключение

Без сомнений, технологии автопилотирования и создания систем управления автомобилями является важным и перспективным направлением, т.к. оно может обеспечить безопасность на дорогах и эффективное использование ресурсов автомобиля.

В данной статье были определены требования к алгоритму, среда разработки, а также используемый инструментарий

Библиографический список

1. Умные автомобили // Центр 2М URL: <https://center2m.ru/ymnie-avtomobili>
2. Беспилотный автотранспорт: Автоматизированные системы управления и информационные технологии. Москва, 2009.
3. Автопилот. Беспилотный автомобиль // TAdviser URL: [https://www.tadviser.ru/index.php/Статья:Автопилот_\(беспилотный_автомобиль\)](https://www.tadviser.ru/index.php/Статья:Автопилот_(беспилотный_автомобиль))
5. Webots // ИТ блог Олега Солозובה URL:<https://8d9.ru/program/webots>
6. Датчики. Актуальные технологии и применения датчиков автомобильных систем активной безопасности//Компоненты и технологии URL: <https://kit-e.ru/sensor/datchiki-chast-5/>
7. ПИД-регулятор. Основные задачи, применение и методика настройки // Elec.ru URL: <https://www.elec.ru/articles/pid-reguljator-metodika-nastrojki/>
8. Уроки Ардуино. ПИД регулятор // youtube URL: <https://www.youtube.com/watch?v=rIbWnB26dp0>
9. Карпов В.Э. ПИД-управление в нестрогом изложении. Москва: 2012.

V.E. Gai, V.F. Zharinov, V.A. Lazareva

SOFTWARE CAR CONTROL SYSTEM IN THE WEBOTS SIMULATOR

Nizhny Novgorod state technical university n.a. R.E. Alekseev

Purpose: Development of a software car control system in the Webots simulator, which implements the safe movement of a car on a multi-lane road with many static and dynamic obstacles.

Design/methodology/approach: open source and multi-platform desktop application «Webots».

Findings: the author described the implementation of a software car control system in the Webots simulator, which includes the safe movement of a car on a multi-lane road with many static and dynamic obstacles.

Originality/value: the study of autopilot technologies has a significant role in the further development of these technologies. This makes it possible to see the advantages and disadvantages of certain techniques and algorithms, and also shows the further potential in their use.

Keywords: autopilot, software control system, PID controller, distance sensors, Webots.

ПРОГРАММНАЯ СИСТЕМА УПРАВЛЕНИЯ КОПТЕРОМ В СИМУЛЯТОРЕ WEBOTS

Нижегородский государственный технический университет им. Р.Е. Алексеева

Цель работы – Снабжение модели коптера возможностью слежения за целью в программной среде Webots. Проведение экспериментального подтверждения его работоспособности.

Ключевые слова: обнаружение объекта, робот, трекинг объекта

Дроны стали настоящим прорывом в технологическом развитии мира. Самыми популярными и более доступными сейчас являются квадрокоптеры. Это летательные аппараты, отличительная особенность которых заключается в свободном полете в трех плоскостях, в том числе и назад. Также данным устройствам не требуются специальные площадки для взлета и посадки, они могут перевозить различные грузы на внешних осях. Сфера применения квадрокоптеров увеличивается с каждым годом. Изначально, это были именно военные устройства, работающие в целях разведки. Один из первых, пилотируемых коптеров с четырьмя роторами, который на практике смог оторваться от поверхности и зависнуть над землей, был создан и испытан русско-американским авиаконструктором Георгием Ботезатом еще в 1922 году. Но на тот момент аппарат еще не могли обеспечить всеми механизмами для стабильной работы, стремительное развитие устройств началось позже.

В настоящее время коптеры используют не только в промышленных, военных, но даже и в развлекательных целях. В теме работы «Программная система управления коптером в симуляторе Webots» основная функция дрона заключается в возможности слежения за целью. Данную функцию можно использовать в военной сфере, в скотоводстве, в фото и видео съемках и просто в сфере развлечений. Дроны сейчас могут работать даже в космосе без участия человека. Студенты, изучающие данную отрасль, в будущем будут иметь больше возможностей для трудоустройств. После создания модели робота в виртуальной среде, проведения испытаний –полученные результаты могут быть перенесены на реальные устройства. Такие симуляторы, как Webots значительно снижают затраты на создание технического средства с различными характеристиками, так как во время работы в программе нам не нужно покупать оборудование и создавать площадки для проведения тестирования.

Программные средства

Реализация проекта будет осуществляться в бесплатном и открытом 3Dсимуляторе роботов Webots [1]. Данная программа обладает большой коллекцией свободно изменяемых моделей роботов, датчиков, исполнительных механизмов и объектов. Она представляет полный набор для программирования и моделирования устройств в различных условиях. Webots использует вилку ODE (Open Dynamics Engine) для обнаружения столкновений и моделирования динамики твердого тела. Библиотека ODE позволяет точно моделировать физические свойства объектов, такие как скорость, инерция и трение.

Сам проект будет состоять из технического средства, которое будет выполнять поставленные задачи, виртуального пространства, в котором будет реализована сцена для проведения испытаний и объекта слежения.

Модель квадрокоптера, выбранная для испытаний - это Mavic 2 PRO (Рис.1).

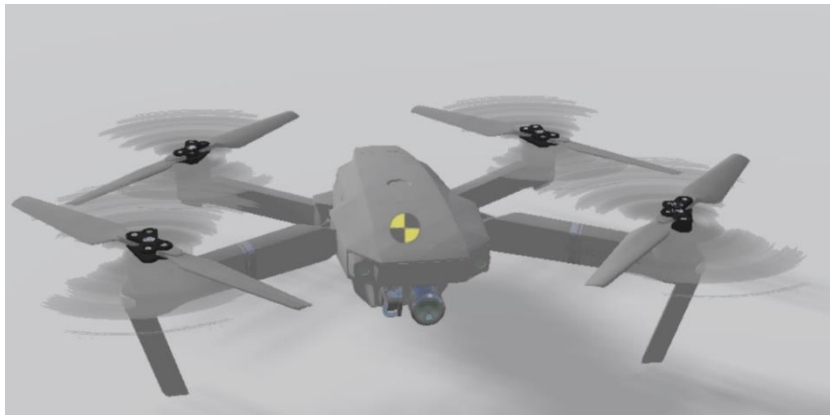


Рис. 1. Модель квадрокоптера - Mavic 2 PRO [5]

Mavic 2 PRO - это беспилотный квадрокоптер от DJI, с установленной моторизованной камерой.

Блок-схема алгоритма

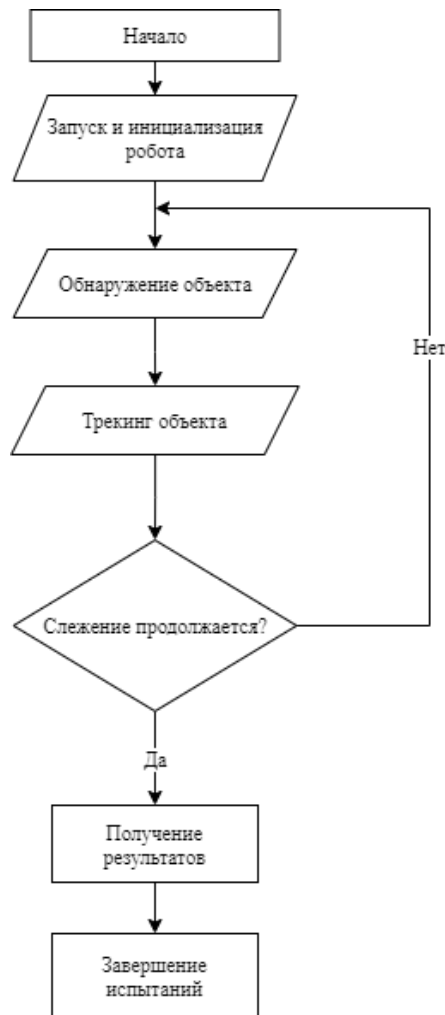


Рис. 2. Блок-схема алгоритма

Требования к алгоритму

Будет создана площадка с различными препятствиями, как находящимися на земле, так и возвышающимися над ней (деревья, дома) и движимый объект. Квадрокоптер с камерой и алгоритмом слежения должен двигаться за объектом, минуя все препятствия, повторяя маршрут объекта, находясь на определенной высоте. Блок – схема представлена на Рис. 2.

Реализация

Для реализации алгоритма системы слежения нужно спрограммировать основной контроллер – программа, обеспечивающая движение дрона по определенным траекториям и в зависимости от ситуации.

Главные компоненты для реализации:

1. PID-регулятор – устройство в управляющем контуре с обратной связью. Служит для контроля движения, совершения маневров. ПИД регулятор берет данные измеренные сенсорами полетного контроллера (гироскопы, акселерометры) и сравнивает их с ожидаемым значениями, чтобы изменить скорость моторов для компенсации любых отклонений и удержания баланса. P – это основное значение, число определяющее стабильность. P определяет, насколько жестко контроллер полета исправляет погрешности, чтобы достичь желаемой траектории полета. Значение I определяет, насколько сильно он будет поддерживать квадрокоптер при воздействии внешних факторов, таких как ветер и смещение центра тяжести, например. Это настройка жесткости при поворотах квадрокоптера. D работает как глушитель и уменьшает чрезмерную коррекцию и регулирование коэффициента P . Увеличивая значение D , происходит смягчение значения P , минимизируется вибрация пропеллеров. PID настраиваются индивидуально, в зависимости от пилота и внешних факторов [2].

2. Детектор объектов – служит для верного обнаружения объекта (Рис.3), за которым должен будет следить коптер, он напрямую связан с компьютерным зрением [4]. Компьютерное зрение состоит из различных аспектов, таких как распознавание изображений, обнаружение объектов, генерация изображений, супер-разрешение изображений и многое другое. Обнаружение объекта ссылается на способность компьютера и систем программного обеспечения находить объекты на изображении/сцене и идентифицировать каждый объект. Обнаружение объектов является наиболее глубоким и сложным аспектом компьютерного зрения в виду огромного количества способов практического применения- обнаружения лиц, обнаружения транспортных средств, подсчета пешеходов, веб-изображений, систем безопасности и машин без водителя. Такое многообразие сфер, в которых может применяться детектор объектов еще раз доказывает актуальность выбранной темы.

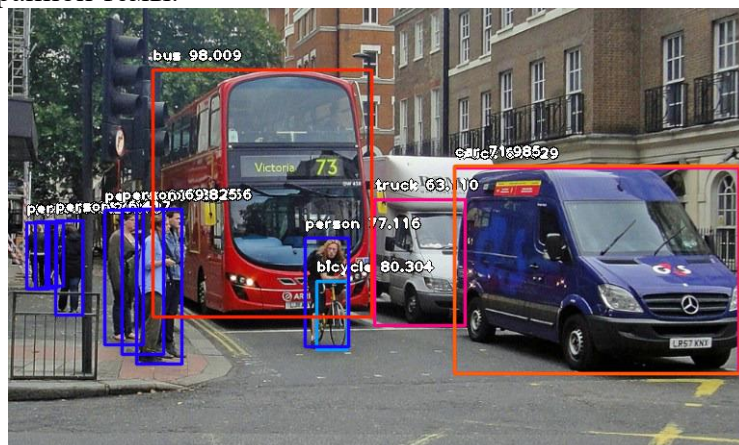


Рис. 3. Изображение после обнаружения объектов

3. Трекер положения объекта. Трекинг - является одним из важных механизмов современных систем видеонаблюдения, позволяющий подсчитывать объекты (люди, автомобили и др.) и вести учёт их перемещений. В моем случае этот трекер как раз будет служить для выполнения основной задачи – слежение за созданным объектом. Вопросы качества работы трекинга упираются в качество идентификации объектов на видео. Первоначально это были HOG-дескрипторы. Затем более качественные особенности — SURF, SIFT, FREAK и др. [3]

На базе новых методов детектирования объектов построены и методы трекинга:

1) ROLO. Метод отслеживания одного объекта, который объединяет обнаружение объектов и рекуррентные нейронные сети. ROLO — это комбинация YOLO и LSTM.

2) Deep SORT. Deep SORT улучшает SORT (алгоритмический подход к отслеживанию объектов), заменяя ассоциированную метрику новым обучением косинусной метрики, методом изучения пространства признаков, в котором косинусное сходство эффективно оптимизируется за счет репараметризации режима softmax.

3) Track R-CNN. Track R-CNN был введен в качестве основы для решения задачи многообъектного отслеживания и сегментации (MOTS). Трекер создается путем интеграции трехмерных сверток.

Заключение

Выбранная тема является актуальной сейчас. Технологии развития различных дронов стремительно растут. Беспилотные устройства заполняют все больше и больше сфер нашей жизни. Изучение данных устройств в университете, пусть и в виртуальном формате, дает больше шансов студентам в трудоустройстве.

В работе изложены основные понятия, сферы и этапы моего исследовательского проекта.

Библиографический список

1. Webots [Электронный ресурс]: Википедия. Свободная энциклопедия. – Режим доступа: <https://en.wikipedia.org/wiki/Webots>
2. ПИД для квадрокоптеров (перевод) [Электронный ресурс]: RCDetails Blog. – Режим доступа: <https://blog.rcdetails.info/pid-dlya-kvadrokopterov-perevod/>
3. Кручинин Александр. Современные технологии трекинга объектов на видео [Электронный ресурс]: Resog.ru. Образовательный сервис – Режим доступа: <http://resog.ru/современные-технологии-трекинга-обь/>
4. DmitryKulbeda. Перевод статьи MosesOlafenwa: ObjectDetectionwith 10 linesofcode [Электронный ресурс]: medium.com. Образовательный сервис. – Режим доступа: <https://medium.com/nuances-of-programming/обнаружение-объектов-с-10-строчками-кода-953bd0e22a2>
6. Webots User Guide [Электронный ресурс]: Информационный портал cyberbotics – Режим доступа: <https://cyberbotics.com/doc/guide/robots>

Gai V.E., Semenova D.S.

SOFTWARE SYSTEM FOR COPTER CONTROL IN THE WEBOTS SIMULATOR

Nizhny Novgorod state technical university. R.E. Alekseeva

The paper describes the process of modeling a software system for controlling a copter in the Webots environment.

Purpose: Providing the copter model with the ability to track the target in the Webots software environment. Conducting an experimental confirmation of its operability.

Main components for implementation: PID controller, object detector and object position tracker.

Features: The model allows you to see all the errors and features in the system. Such research minimizes costs, since there is no need to assemble the device itself

Software tools: Webots simulator, in which the copter will be endowed with the ability to track the target

АППАРАТНО - ПРОГРАММНАЯ СИСТЕМА, РЕАЛЕЗУЮЩАЯ ИГРУ РАСМАН

Нижегородский государственный технический университет им. Р.Е. Алексеева

Задача этой работы заключается в разработке и создании аркадного автомата на базе RaspberryPi в корпусе типа «Upright» и написание под этот автомат простой игры с искусственным интеллектом на основе аркады PacMan.

Ключевые слова: Аппаратно-программная система, искусственный интеллект, игры.

В современном мире игровая индустрия развивается семимильными шагами, выходит множество новых игр и с каждым днем людей, увлеченных видео играми, становится все больше. Проходят турниры по многим дисциплинам с огромными призовыми и крупными спонсорами, да и киберспорт уже давно признан официальным видом спорта. При всем разнообразии современных игр с реалистичной графикой и красочными открытыми мирами, большое количество людей все еще поигрывают время от времени в старые аркадные игры с которых начиналась вся эта огромная индустрия. Я тоже захотел погрузиться в мир старых игр, но просто скачать эмулятор на ПК будет слишком скучно, поэтому собрать свой аркадный автомат и написать под него простенькую игру с искусственным интеллектом показалось мне хорошей идеей.

Аппаратное обеспечение

Данный аркадный автомат разрабатывается на платформе Raspberry Pi 3B+ [1]. Для подобного аппарата понадобится: сама RaspberryPi3B+, динамики для вывода звука (можно взять любые старые колонки), блок питания от RaspberryPi (обычно идет в комплекте), любой монитор с выходом HDMI, набор кнопок и джойстиков.

У аркадных автоматов существует несколько стандартных типов корпусов.

Вертикальный корпус (Upright)–Это самый распространённый тип корпуса. Обычно выполнен из дерева или металла с высотой около двух метров и шириной 60 сантиметров. Дисплей расположен на уровне головы игрока, а игровая панель в районе пояса.

Горизонтальный корпус – этот вид корпуса напоминает невысокий стол, по разным сторонам которого расположены элементы управления, а по середине стола установлен монитор. В играх для нескольких игроков монитор выводит изображение для каждого игрока в свой ход, когда происходит передача управления другому игроку то изображение на мониторе просто переворачивается в сторону данного игрока.

Сидячий корпус – наиболее распространен в Японии. Чаще всего сделан из пластика и металла, по высоте немного ниже чем вертикальных тип корпуса. Зачастую оборудован большим монитором. Игрок располагается перед автоматом на стуле.

В нашем проекте был выбран вертикальный тип корпуса, так как он наиболее популярен и привычен для большинства людей. И при создании такого типа корпуса будет максимально сохранена атмосфера ностальгии, к которой мы и стремимся.

Корпус для аппаратного обеспечения планируется собрать из мебельного ЛДСП. Крепление – обычные строительные уголки, покраска автомата также осуществляется своими силами с использованием обычной краски, которую можно найти в любом магазине. Размеры для корпуса были позаимствованы в интернете [2].

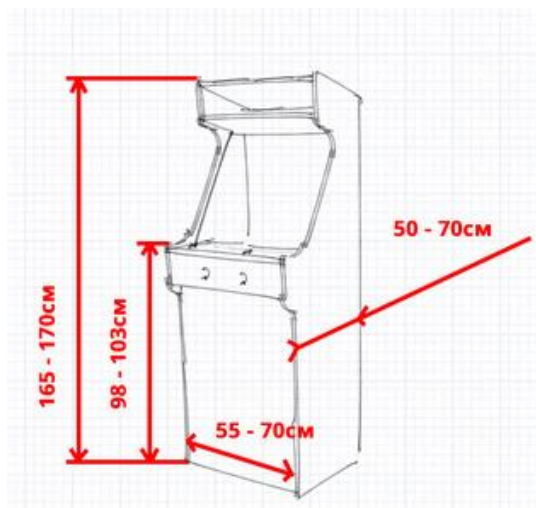


Рис. 1. Примерные размеры корпуса

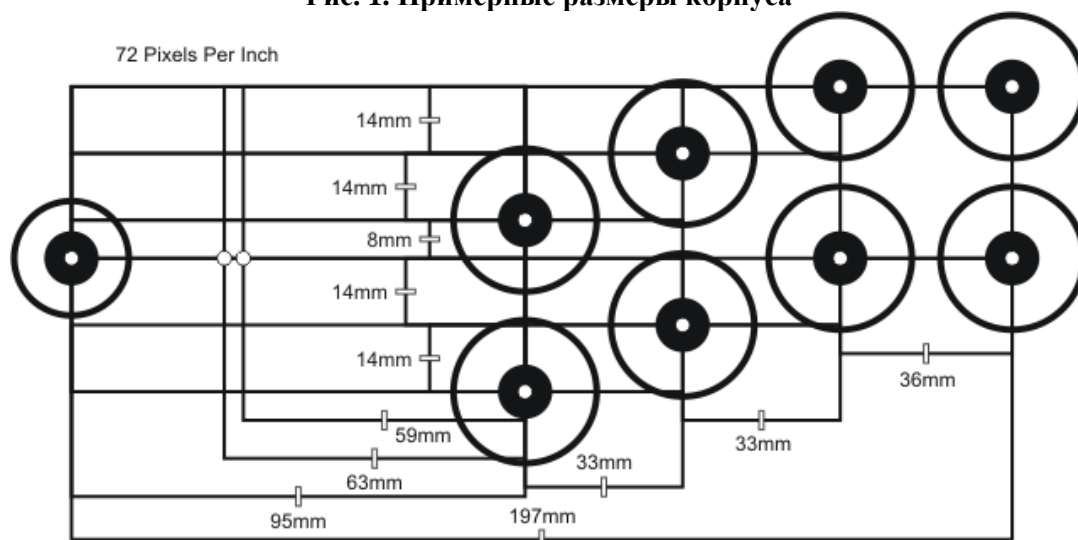


Рис. 2. Схема расположения джойстика и кнопок

Подключение внутренних компонентов выполняется следующим образом:

- разъем USB-контроллера подключается к USB разъёму мини-ПК;
- HDMI выход на RaspberryPi подключается к входу HDMI платы контроллера;
- динамик припаивается к плате усилителя;

Питание всех элементов управления происходит через RaspberryPi, монитор питается от розетки 220В через удлинитель, который спрятан в корпусе. Далее идет установка программного обеспечения.

Программное обеспечение

В качестве программного обеспечения для данного аркадного аппарата было решено написать простую игру с искусственным интеллектом. Так как придумать новую игру не просто, то было решено взять за основу всем известную игру PacMan. Также если взять за основу старую уже всем знакомую игру, то сохранится атмосфера ностальгии, для которой обычно и покупают аркадные аппараты.

Pac-man – это аркадная игра, разработанная в 1980 году японской компанией Namco. Смысл игры заключается в том, чтобы пройти по лабиринту съев при этом все точки, раскиданные по нему, и не попасть на приведения, которые бегают по лабиринту в погоне за героем. В игре существуют разные уровни сложности, уровень повышается если игрок прошел предыдущий, при этом карта на протяжении всех уровней остается неизменной. На каждом уровне в PacMan есть несколько дополнительных жизней, которые

тратятся если до героя дотронулся призрак. Когда происходит столкновение то герой и все призраки расставляются по своим изначальным местам, но точки, которые уже удалось съесть не восстанавливаются, и так пока не закончатся все жизни у героя.

Всего в лабиринте есть 240 маленьких точек и 4 большие точки. За съедание героем маленькой точки игроку дается 10 очков, за съедание большой точки 50 очков. Таким образом, в общей сложности все точки в лабиринте стоят 2600 очков. На ранних уровнях при съедании героем большой точки все призраки в лабиринте на короткое время входят в режим испуга, в этом режиме призраки резко начинают движение в другую сторону, и герой имеет возможность их съесть, за съедание призрака также даются дополнительные очки. После того как призрак был съеден, он возвращается на поле из центра лабиринта.



Рис. 3. Вид оригинальной игры Pac-Man

Всего в PacMan 4 разных призрака, желтый, синий, красный и розовый и они имеют три разные запрограммированные модели поведения: преследование, рассеивание и испуг. Эти модели поведения присущи всем призракам. Преследование считается режимом по умолчанию, рассеивание иногда включается на короткий промежуток времени, а вот режим испуг активируется на короткое время при поедании героем большой точки. Также у каждого призрака существуют как общие черты поведения, так и свои индивидуальные особенности.

Рассмотрим общие черты их поведения подробнее. Искусственный интеллект призраков [3] очень прост и недалек, они думают лишь на один шаг в будущее, когда движутся в лабиринте. Всякий раз, когда призрак ступает на клетку он смотрит вперед на следующую и смотрит куда он повернет, когда достигнет ее. Призраки никогда не могут изменить направление движения на противоположное, то есть призрак не может войти на клетку с левой стороны, а затем решить начать двигаться в обратном направлении и выйти слева. Смысл этого ограничения в том, что игра всегда будет продолжать идти в том же направлении, так как когда призрак входит на клетку лишь с двумя выходами, он не может выбрать тот, откуда пришел. Единственное исключение из правила, это смена режима с преследования или разбегания на любой другой. В этом случае приведения меняют свое направление, что обязательно отменяет любое решение о предыдущем направлении движения.

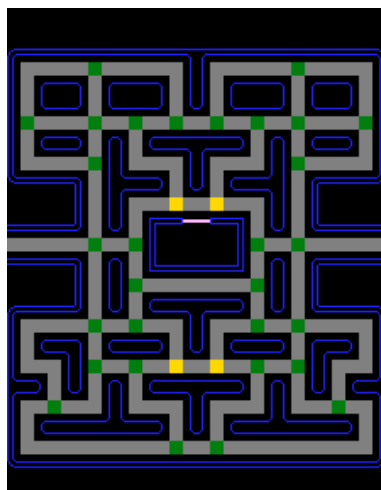


Рис. 4. Упрощенная схема лабиринта

Зеленым обозначены клетки, когда необходимо принять решение, тогда призрак выбирает прилегающую клетку, которая поставит его ближе по прямой линии к цели. Так как расстояние выбирается по прямой до цели, возможно, что призрак выберет неправильный поворот, когда расстояние по прямой меньше, однако расстояние по лабиринту может оказаться больше.

Клетки, обозначенные желтым цветом, имеют дополнительные ограничения, призраки попав на эти клетки не могут пойти вверх. Придя на них слева или справа, призраки могут только пойти дальше, исключением являются случаи при смене режима. Однако в режиме испуга призраки двигаются псевдослучайно на каждом повороте, и это ограничение в таком случае не работает.

Заключение

Результатом данной работы является разработанный и собранный аркадный аппарат на базе RaspberryPi, предназначенный под различные старые игры. Также написана игра на основе всем известной аркады Pac-Man использованием искусственного интеллекта.

В рамках данной статьи рассмотрены основные элементы аркадного автомата, в частности элементы управления, материалы и размеры корпуса, также рассмотрены основные этапы сборки аппарата. Рассмотрен принцип игры, на которой будет работать разработанный аркадный аппарат.

Библиографический список

1. Знакомство с RaspberryPi //Amperka: сайт. – 2020. - URL:<http://wiki.amperka.ru/articles:rpi-guide> (дата обращения: 15.01.2021).
2. PanelLayout // Slagcoin: сайт. – 2008-2009. - URL:<http://www.slagcoin.com/joystick/layout.html>(дата обращения: 15.01.2021).
3. Алгоритм поведения привидений в игре Pac-Man // Habr: сайт. – 2010. -URL: <https://habr.com/ru/post/109406/>(дата обращения: 15.01.2021).
4. Алекс Дж. Шампандар. Искусственный интеллект в компьютерных играх // Издательский дом «Вильямс», 2007. – 768с.

A. A. Polyakov, V. E. Gai

**HARDWARE AND SOFTWARE SYSTEM THAT IMPLEMENTS
THE PACMAN GAME**

Nizhny Novgorod state technical university named after R. E. Alekseev

Purpose: development and creation of an arcade machine based on Raspberry Pi in the "April" type case and writing a simple game with artificial intelligence based on the PacMan arcade for this machine.

Design/methodology / approach: assembly, programming, artificial intelligence.

Findings: the result is a developed and assembled arcade machine based on the Raspberry Pi, with a written game based on the arcade Pac-Man using artificial intelligence.

Research limitations/implications: the arcade machine can be used for leisure activities.

Keywords: hardware and software system, artificial intelligence, games.

ПРОГРАММНАЯ СИСТЕМА ФОРМИРОВАНИЯ КАРТЫ ПОМЕЩЕНИЯ НА ОСНОВЕ ОБУЧЕНИЯ С ПОДКРЕПЛЕНИЕМ

Нижегородский государственный технический университет им. Р.Е. Алексеева

Задача состоит в том, чтобы робот по мере маршрута изучал местность, огибая все препятствия. Актуальность данной темы возникает в случае исследования расположения объектов в помещении и самого помещения в целом. Возможно выявление аномального расположения объекта, а также расположение странных объектов (тело человека, животного и т. п.), которые могут помочь в выявление некоторых проблем.

Ключевые слова: обучение с подкреплением, агент, вознаграждение, PID – регулятор, робот, карта помещения.

Обучение с подкреплением – обучение посредством взаимодействия с средой (моделью). В основном основана на принципе поощрений [1, 2].

Существует несколько подходов обучения с подкреплением:

1. Основанная на модели. Изначально дана модель со всеми вознаграждениями, и мы по ней учим робота выбирать оптимальный путь;
2. Обучение по значениям. Имеется шпаргалка с вознаграждениями на следующем шаге, и мы выбираем наш путь, учитывая известные шаги и сумму их (какой исход лучше по тому и пойдём);
3. Q-обучение. Для каждого шага формируется полезность. После уже агент выбирает свой путь не случайно, а по уже сформированной стратегии предыдущего опыта;
4. Policy-Gradient. Позволяет оптимизировать стратегию агента относительно случайности окружающей среды.

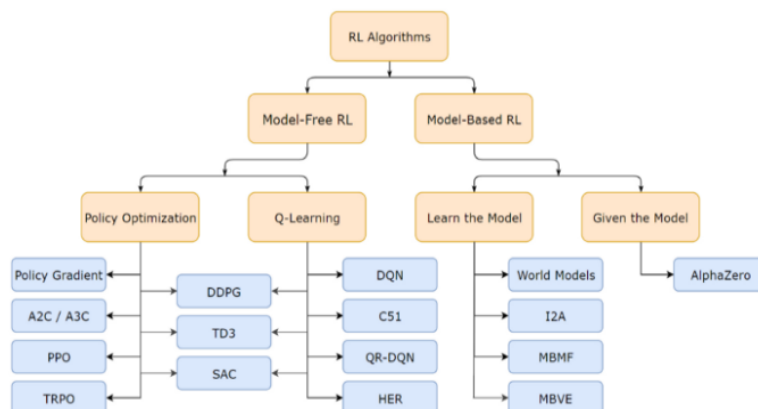


Рис. 1. Классификация алгоритмов обучения с подкреплением

В данной задаче можно использовать метод Policy-Gradient, ведь для передвижения и расчетов нужно будет учитывать препятствия, которые не будут иметь четкого расположения, а будут разбросаны случайно.

Алгоритм Policy-Gradient

Вероятность реализации сценария зависит от поведения среды, которое задается вероятностями перехода между состояниями $p(s_{t+1} | s_t, a_t)$, распределением начальных состояний $p(s_1)$ и поведения агента, которое определяется его стохастической стратегией $\pi_\theta(a_t | s_t)$

$$p_{\theta}(\tau) = p_{\theta}(s_1, a_1, \dots, s_T, a_T) = p(s_1) \prod_{t=1}^T \pi_{\theta}(a_t | s_t) p(s_{t+1} | s_t, a_t).$$

Вероятности переходов между состояниями агенту неизвестны, то есть у агента нет модели поведения окружающей среды. Нужно выбрать такой набор параметров агента θ , задающий $\pi_{\theta}(a | s)$, чтобы максимизировать мат ожидание суммы полученных выигрышей:

$$J(\theta) = E_{\tau \sim p_{\theta}(\tau)} [R_{\tau}] = \int p_{\theta}(\tau) R_{\tau} d\tau,$$

$$p_{\theta}(\tau) \nabla_{\theta} \log p_{\theta}(\tau) = p_{\theta}(\tau) \frac{\nabla_{\theta} p_{\theta}(\tau)}{p_{\theta}(\tau)} = \nabla_{\theta} p_{\theta}(\tau),$$

$$\nabla_{\theta} J(\theta) = \int \nabla_{\theta} p_{\theta}(\tau) R_{\tau} d\tau,$$

$$\nabla_{\theta} J(\theta) = E_{\tau \sim p_{\theta}(\tau)} [\sum_{t=1}^T \nabla_{\theta} \log \pi_{\theta}(a_t | s_t) R_{\tau}].$$

В получившееся выражение для $\nabla_{\theta} J(\theta)$ уже не входят напрямую значения $p(s_{t+1} | s_t, a_t)$ и $p(s_1)$, которые неизвестны. Таким образом, если есть в наличии сценарий τ и соответствующее ему значение R_{τ} , можно вычислить величину $\sum_{i=1}^N (\sum_{t=1}^T \nabla_{\theta} \log \pi_{\theta}(a_t | s_t^i)) R_{\tau}$. Значит, если есть выборка из N уже известных сценариев $\tau^i = (s_1^i, a_1^i, \dots, s_T^i, a_T^i)$, полученная из распределения $\tau \sim p_{\theta}(\tau)$, то можно приблизить посчитать приблизительное значение $\nabla_{\theta} J(\theta)$ по методу Монте-Карло — вычислив выборочное среднее случайной величины:

$$\nabla_{\theta} J(\theta) \approx \frac{1}{N} \sum_{i=1}^N (\sum_{t=1}^T \nabla_{\theta} \log \pi_{\theta}(a_t^i | s_t^i)) R_{\tau^i} = \frac{1}{N} \sum_{i=1}^N (\sum_{t=1}^T \nabla_{\theta} \log \pi_{\theta}(a_t^i | s_t^i)) (\sum_{t=1}^T r(s_t^i | a_t^i)).$$

Таким образом, оптимизировать $J(\theta)$ можно с помощью следующего простого алгоритма:

1) Прогнать N сценариев τ_i со стратегией $\pi_{\theta}(a | s)$

2) Посчитать среднее

арифметическое $\nabla_{\theta} J(\theta) \leftarrow \frac{1}{N} \sum_{i=1}^N (\sum_{t=1}^T \nabla_{\theta} \log \pi_{\theta}(a_t^i | s_t^i)) (\sum_{t=1}^T r(s_t^i | a_t^i))$

3) $\theta \leftarrow \theta + \alpha \nabla_{\theta} J(\theta)$

4) Если не сошлись к экстремуму, повторить с пункта 1

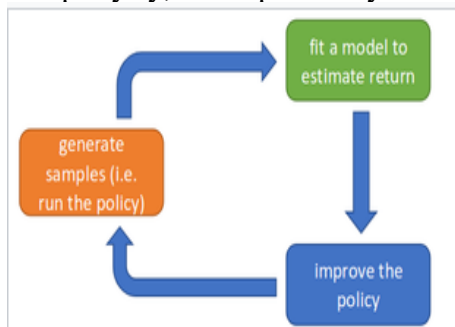


Рис. 2. Policy Gradient

PID – Регулятор

PID – Регулятор представляет собой совокупность 3 компонентов, а именно пропорциональный, интегрирующий и дифференцирующий компоненты. Он используется для обеспечения точности передаваемых ему значений [3].

Компоненты PID – регулятора [3]:

1) Пропорциональная $P(t) = K_p * e(t)$. Работает по принципу: чем больше отклонение, тем большее значение будет сформировано

2) Интегрирующая $I(t) = K_i * \int_0^t e(t)dt$. Выявляет статические ошибки

3) Дифференцирующая $D(t) = K_d * \frac{de(t)}{dt}$. Выявляет отклонения возможные в будущем, учитывая действующие величины.

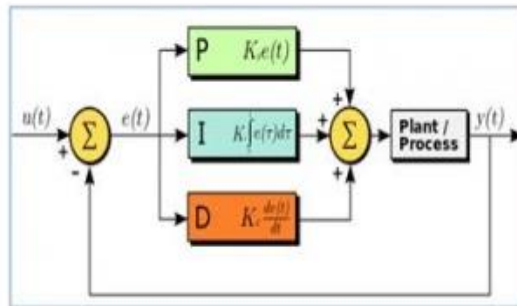


Рис. 3. PID-регулятор

Для робота понадобится PID – регулятор, ведь регулировка скорости и траектории движения должна будет производиться должным образом.

Алгоритм

Робот будет двигаться по помещению, выбирая выгодный для себя маршрут. По мере движения робот будет огибать препятствия (в данном случае обучение с подкреплением научит робота выбирать маршрут без препятствий на пути), а также с помощью датчиков будет рисовать ту часть помещения, которую он проезжает, для того, чтобы можно было потом увидеть размер помещения и расположения объектов на нем.

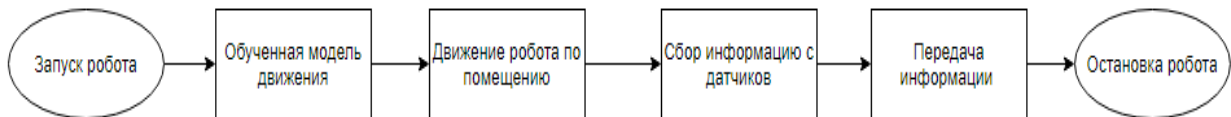


Рис. 4. Алгоритм программы

Программное обеспечение

В языках программирования Python и Java имеются готовые библиотеки с реализованным обучением с подкреплением

1) Java:

a. Teachingbox - инструментарий для разработки алгоритмов обучения с подкреплением

b. BURLAP (Brown-UMBC Reinforcement Learning and Planning) — библиотека алгоритмов планирования агентов и обучения с подкреплением

2) Python

a. MMLF (Maja Machine Learning Framework) — библиотека алгоритмов обучения с подкреплением и набор тестовых сред для их проверки

b. OpenAIGym - платформа для разработки и сравнения алгоритмов обучения с подкреплением от OpenAI

c. PyBrain - библиотека алгоритмов машинного обучения

d. RLPY - библиотека для проведения экспериментов по обучению с подкреплением

Среди представленных инструментов OpenAIGym подходит, так как он довольно прост. Так же для этого будет использоваться Python 3.7. Так же в Python уже имеется библиотека с готовым PID – регулятором.

Заклучение

В данной работе было рассмотрено обучение с подкреплением, его алгоритмы, а также был произведен выбор алгоритма для нашей задачи с движением робота. Был рассмотрен принцип работы PID – регулятора.

Библиографический список

1. Kaelbling Leslie P.; Littman Michael L.; Moore Andrew W. (1996). "Reinforcement Learning: A Survey"
2. van Otterlo M.; Wiering M. (2012). Reinforcement learning and markov decision processes. Reinforcement Learning. Adaptation, Learning, and Optimization. 12. pp. 3–42. doi:10.1007/978-3-642-27645-3_1. ISBN 978-3-642-27644-6.
3. Kiam Heong Ang; Chong G.; Yun Li (2005). "PID control system analysis, design, and technology" (PDF). IEEE Transactions on Control Systems Technology. 13 (4): pp.559–576. doi:10.1109/TCST.2005.847331. S2CID 921620.
4. Gosavi Abhijit (2003). Simulation-based Optimization: Parametric Optimization Techniques and Reinforcement. Operations Research/Computer Science Interfaces Series. Springer. ISBN 978-1-4020-7454-7.
5. Burnetas Apostolos N.; Katehakis Michael N. (1997), "Optimal adaptive policies for Markov Decision Processes", Mathematics of Operations Research, 22: 222–255, doi:10.1287/moor.22.1.222
7. Williams Ronald J. "Simple statistical gradient-following algorithms for connectionist reinforcement learning." Machine learning 8.3-4 (1992): 229-256.
8. Sutton Richard S., et al. "Policy gradient methods for reinforcement learning with function approximation." Advances in neural information processing systems. 2000.
9. Policy Gradients. CS 294-112: Deep Reinforcement Learning, Sergey Levine.

V.A. Popov V.E. Gai, V.P. Zelensky

SOFTWARE SYSTEM FOR FORMING A ROOM MAP BASED ON TRAINING WITH REINFORCEMENT

Nizhny Novgorod state technical university. n.a. R.E. Alekseev

The paper describes the process of creating A software system for generating a room map based on reinforcement learning

Purpose: learning terrain in the all way around all obstacle.

Design/methodology/approach: machine learning algorithms, reinforcement learning.

Findings: the author was learned robot explore room and generate room map

Research limitations/implications: reinforcement learning takes a lot of time for fool learning and fool learning not equal 100 percent

Originality/value: this task come up in case of research a position of all object room and room on the whole. Potentially detection the abnormal object or unusual object (human body, animal body or etc.), which can help in detection some problems.

РАСПОЗНАВАНИЕ ЭМОЦИЙ ЧЕЛОВЕКА ПО ИЗОБРАЖЕНИЮ ЛИЦА

Нижегородский государственный технический университет им. Р.Е. Алексеева

В работе описывается процесс создания приложения для распознавания эмоций человека по изображению лица с использованием алгоритмов глубокого обучения.

Ключевые слова: распознавание эмоций человека, эмоции, машинное обучение, МО, свёрточная нейронная сеть, задача классификации, VGG16, CNN, DNN, MTCNN, FER, ML.

В текущий момент крупный бизнес заинтересован в получении обратной связи от своих клиентов. Чаще всего это выражается в написании отзыва на сайте компании, оценке по смс или нажатии соответствующего смайлика на пульте около окна обслуживания клиента. Такие способы не являются абсолютно достоверными, поскольку есть возможность подделать отзыв. Чтобы избежать подобной ситуации можно оценивать эмоции человека вовремя и после обслуживания и получить более достоверную информацию.

Технология распознавания эмоций может быть полезна в любой сфере деятельности, где идёт работа с клиентами в живую, либо удалённо с использованием видеосвязи. Эта технология может помочь понять, насколько человек удовлетворён предоставленными ему услугами.

Отходя в сторону от эмоций, подобная система может пригодиться в медицине. Резкая боль, предынсультное состояние сильно отражается на выражении лица человека и сопровождается определённой мимикой, система, способная детектировать подобные изменения может быть очень полезна, и, в комплексе с другими системами детектирования и оповещения, возможно, спасёт человеку жизнь.

Постановка задачи

Постановка задачи включает в себя следующие этапы:

1. Выделение основных эмоций.
2. Сбор данных. Поиск и объединение необходимых наборов данных
3. Выбор инструментария разработки
4. Предобработка данных
5. Обучение нейронной сети
6. Практическая реализация

Выделение основных эмоций

Эмоция- выраженная мимикой, позой, жестом, вербально, реакция человека на внешний или внутренний раздражитель. Также можно сказать, что эмоция – выражение чувств и состояния человека на данный момент [1].

Существует огромное множество эмоций: гнев, страх, ярость, радость, веселье, интерес и прочее, однако, как правило выделяют 7 базовых эмоций: радость, грусть, гнев, удивление, отвращение, страх, нейтральная. Остальные эмоции можно принимать, как варианты базовых Рис.1.



Рис. 1. Спектр эмоций по выраженности и окраске

Сбор данных. Поиск и объединение необходимых наборов данных

Для обучения нейронной сети необходимо найти подходящие наборы данных. Существует несколько критериев выбора набора данных:

➤ Доступность:

- ◆ открытые бесплатные (Такие наборы данных доступны для скачивания любым пользователям. Получить такие наборы проще всего, однако в них может быть много ошибок, поскольку требования к качеству-минимальны.)
- ◆ открытые по запросу (Такие наборы данных, чаще всего, доступны студентам, и сотрудникам университетов, поскольку для получения доступа к набору данных нужно указать должность и университет.)
- ◆ платные по запросу (Такие наборы данных чаще всего используются небольшими компаниями, не имеющими возможность самостоятельно создать подобный набор данных.)
- ◆ закрытые (Закрытые наборы данных сформированы одной компанией или группой компаний. Такие наборы данных недоступны никому за пределами компании.)

➤ Объём:

- ◆ <5 тысяч – относительно небольшой набор данных.
- ◆ <50 тысяч – нормальный размер набора данных.
- ◆ <300 тысяч – большой набор данных.
- ◆ >300 тысяч – очень большой набор данных.

➤ Качество подготовки данных:

- ◆ Подготовленные – данные, собранные для конкретной задачи, очищенные от пустых значений и выбросов, нормализованные, т.е. имеющие стандартизированный диапазон значений независимых переменных или признаков данных (сведение к интервалам $[0, 1]$ или $[-1, +1]$), готовые к использованию.
- ◆ Неподготовленные – данные требующие очистки и нормализации. Такие данные не готовы к непосредственному использованию и требуют дополнительных манипуляций с ними.

➤ Формат набора данных:

- ◆ Документ содержащий ссылки на изображения
- ◆ Документ содержащий изображения в виде массива байтов
- ◆ Структурированная директория с изображениями

Для текущей задачи был выбран популярный набор данных FER 2013- это открытый бесплатный набор данных в формате .csv, содержащий 34034 уникальных записей. Набор содержит следующие признаки:

- *Emotion* – (integer [0;6]) данный признак представлен целочисленными значениями от 0 до 6. Цифры являются кодом эмоции.
- *Pixels* – (arrayofintegers) изображение 48x48 представленное массивом байт.
- *Usage* – (string [training; public test; other]) признак показывает назначение данной записи: использование в тестовой (10%) или обучающей (80%) выборке. Также присутствует значение Other. Его можно использовать в той или иной выборке.

Выбор инструментария разработки

Для данной работы в качестве инструмента разработки был выбран язык программирования Python 3, поскольку он прост в освоении, имеет широкий набор библиотек, а также легко читать программу, написанную на Python.

Для создания нейросети используются библиотеки TensorFlow, OpenCV, Keras, Matplotlib, NumPy.

В качестве среды разработки выбор пал на платформу GoogleColaboratory. Её использование бесплатно. Это облачный сервис, направленный на упрощение исследований в области машинного и глубокого обучения. Используя Colaboratory, можно получить удалённый доступ к машине с подключенной видеокартой, что позволит снять нагрузку с пользовательского ПК. Также в GoogleColaboratory уже установлено большинство библиотек, необходимых для написания нейросетей, и при необходимости есть возможность самостоятельно легко установить другие библиотеки.

Предобработка данных

Поскольку исходные изображения имеют сложный фон, различные размеры, различные оттенки и другие факторы, необходимо выполнить ряд процессов предварительной обработки изображений, прежде чем выражения лица будут введены в сеть для обучения. Во-первых, необходимо найти лицо на изображении и выделить изображение лица. Во-вторых, нормализовать изображение лица до определенного размера. В-третьих, выровнять гистограмму изображения, чтобы уменьшить влияние освещенности и других факторов.

Первым этапом предобработки изображения является обнаружение и выделение лица. Входное изображение может быть любого разрешения, на нем может быть не одно лицо, либо лицо может занимать лишь небольшой участок изображения. Для дальнейшей работы с этим изображением необходимо найти и вырезать лицо человека, и уже с ним производить работу.

Из нейросетевых подходов в детектировании лиц особенно значимым является Multi-task Cascaded CNN (MTCNN).

Идея MTCNN — использовать для предсказания положения лица и его особых точек три нейросети последовательно. Особых точек лица в данном случае ровно 5: левый глаз, правый глаз, левый край губ, правый край губ и нос. Первая нейросеть из каскада (P-Net) используется для генерации потенциальных регионов лица. Вторая (R-Net) — для улучшения координат полученных боксов. Третья (O-Net) нейросеть ещё раз регрессирует координаты боксов и, помимо того, предсказывает 5 ключевых точек лица. Multi-task эта сеть потому, что решаются три задачи: регрессия точек боксов, классификация лицо/не лицо для каждого бокса и регрессия точек лица. Причём MTCNN делает это всё в real-time, то есть ей требуется менее 40 ms на одну картинку[2]. На Рис.2 представлены части MTCNN - модели P-Net, R-Net, O-Net.

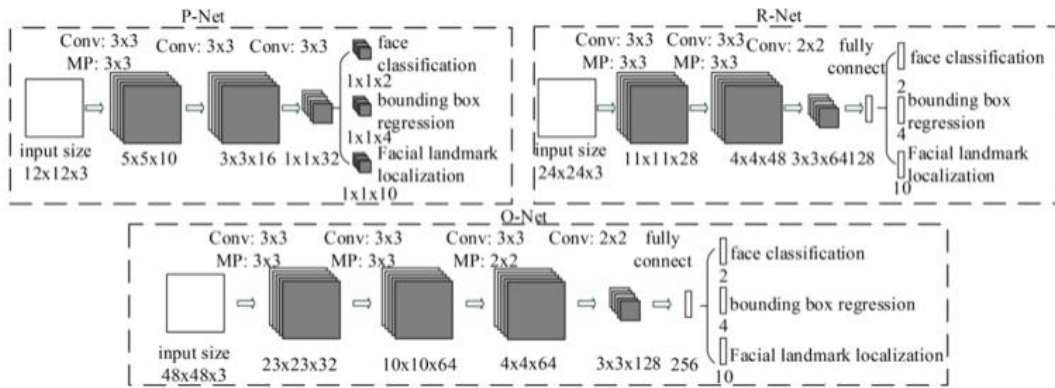


Рис. 2. Составные части MTCNN. P-Net, R-Net, O-Net

Второй этап- масштабирование изображения. Все изображения, поступающие на входной слой нейросети должны иметь одинаковое разрешение. Не рекомендуется использовать большое разрешение, т.к. обучение может стать очень долгим. Было выбрано разрешение 48x48- такие изображения быстро обрабатываются нейросетью, при этом разрешение достаточное для определения эмоции.

На третьем этапе изображение из цветного преобразуется в серое, а также повышается контрастность изображения. Все преобразования над изображением могут быть выполнены с применением библиотеки OpenCV.

Обучение нейронной сети

Для работы была выбрана модель свёрточной нейронной сети VGG16.

Сверточная нейронная сеть (Convolutional Neural Network — ConvNet/CNN) — это Deep Learning-алгоритм, который может принимать входное изображение, присваивать важность (усваиваемые веса и смещения) различным областям/объектам в изображении и может отличать одно от другого. CNN применяется для задач классификации, детектирования, распознавания.

Сеть состоит из большого количества слоёв. После начального слоя (входного изображения) сигнал проходит серию свёрточных слоёв, в которых чередуется собственно свёртка и пулинг). Чередование слоёв позволяет составлять «карты признаков» из карт признаков, на каждом следующем слое карта уменьшается в размере, но увеличивается количество каналов. На практике это означает способность распознавания сложных иерархий признаков.

CNN может состоять из неограниченно большого количества слоёв, но обязательными должны быть: входной слой, слой свёртки, слой объединения (пулинг), полно связный слой, выходной слой.

Inputlayer. Входной слой- слой на вход которого подаётся изображение определённого размера (48x48 в моём случае)

Convolutionallayer. Слой свёртки — это основной блок свёрточной нейронной сети. Слой свёртки включает в себя для каждого канала свой фильтр, *ядро свёртки* которого обрабатывает предыдущий слой по фрагментам (суммируя результаты поэлементного произведения для каждого фрагмента). Весовые коэффициенты ядра свёртки (небольшой матрицы) неизвестны и устанавливаются в процессе обучения.

Poolinglayer. Слой пулинга представляет собой нелинейное уплотнение карты признаков, при этом группа пикселей (обычно размера 2x2) уплотняется до одного пикселя, проходя нелинейное преобразование.

Преобразования затрагивают непересекающиеся прямоугольники или квадраты, каждый из которых ужимается в один пиксель, при этом выбирается пиксель, имеющий максимальное значение. Операция пулинга позволяет существенно уменьшить пространственный объём изображения.

Full-connectedlayer. Вход полно связного слоя должен быть одномерным массивом, тогда как выход предыдущего объединительного слоя S2 - двумерным массивом. Сначала двумерный массив, соответствующий каждому графу признаков, преобразуется в одномерный массив, а затем 128 одномерных массивов последовательно соединяются с вектором признаков размером 51200 ($20 \times 20 \times 128 = 51200$) в качестве полной связи.

Последний слой CNN использует классификатор Softmax. Классификатор Softmax - это конкурентный классификатор с несколькими выходами. Когда заданная выборка вводится, каждый нейрон выводит значение от 0 до 1, которое представляет вероятность того, что входная выборка принадлежит к этому классу. Поэтому в качестве результата классификации выбирается категория, соответствующая нейрону с наибольшим выходным значением [3].

Выходной слой состоит из 7 нейронов, по одному на каждую эмоцию.

От архитектуры нейронной сети многое зависит. Одной из лучших архитектур свёрточной нейронной сети для распознавания объектов на изображении и выделения признаков является VGG16 [4].

VGG16 - модель свёрточной нейронной сети.

Для увеличения точности распознавания эмоций число эпох обучения было доведено до 250.

Но улучшение нейросети при увеличении числа эпох происходит не всегда. Часто происходит проблема, при которой нейросеть отлично работает на уже известных ей изображениях, но на других она теряет свою точность. Это означает проблему переобучения.

Для борьбы с этой проблемой были использованы следующие методы: Dropout и Earlystoppingcallback

Dropout (“метод прореживания”, “метод исключения”). Главная идея Dropout - вместо обучения одной DNN(DeepNeuralNetwork) обучить ансамбль нескольких DNN, а затем усреднить полученные результаты [5].

Сети для обучения получают с помощью исключения из сети (droppingout) нейронов с вероятностью p . На Рис.3 показан принцип работы Dropout.

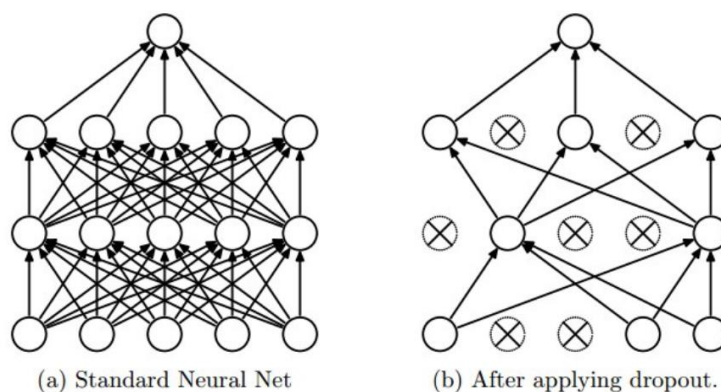


Рис. 3. Принцип работы метода Dropout

Earlystoppingcallback. Чтобы взаимодействовать с нейросетью во время её обучения существует *callback* (Калбэк). С помощью этого инструмента можно отслеживать различные параметры в процессе обучения. На протяжении всего обучения точность распознавания на тренировочных данных будет расти, но точность на проверочных (validationdata) данных будет расти до какого-то определённого показателя. Как только эта точность перестаёт расти- происходит переобучение, следовательно, следует остановить обучение.

Практическая реализация

На вход программы подаётся цветное изображение произвольного размера. На изображении должно быть обязательно видно лицо человека, при этом оно может иметь небольшой поворот или наклон головы. Результат работы программы можно увидеть на Рис.4.

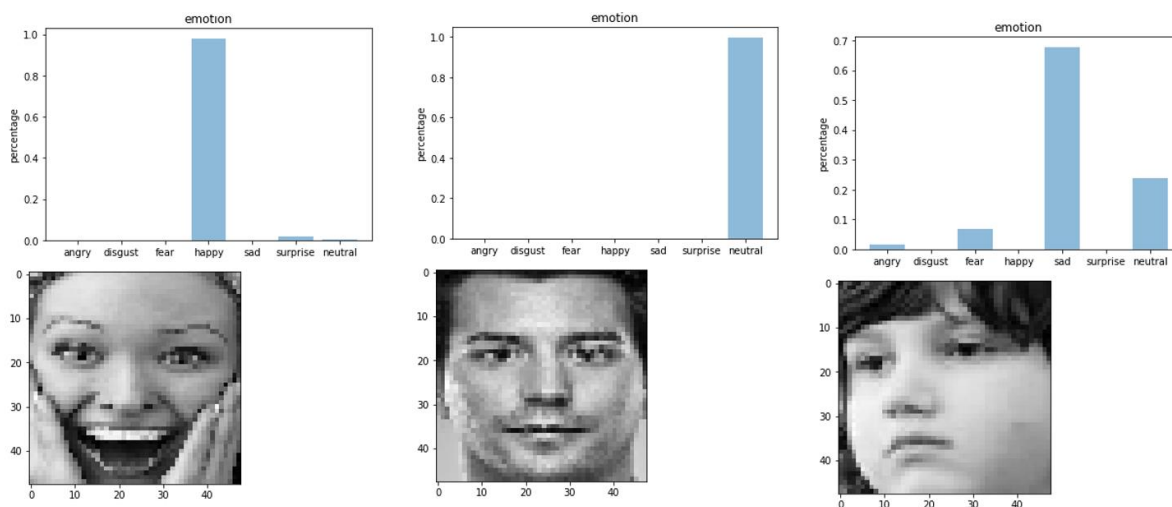


Рис. 4. Результат работы программы

Заключение

Таким образом, на данном этапе работы, была разработана система, позволяющая распознавать эмоции человека по изображению лица. В качестве модели нейронной сети была использована сеть VGG16 с небольшими изменениями. Для обучения был использован набор данных FER2013 содержащий 34 000 записей. Точность алгоритма составляет около 70%, в качестве одного из направлений дальнейшей работы планируется увеличить точность за счёт модификации модели сети, добавлении ещё одного набора данных, а также применения GAN.

В рамках данной работы была создана нейросеть и минимальный интерфейс для загрузки изображений. В планах- создание GUI, и перенос нейросети в программу на Python3.

Библиографический список

1. Экман П. “Эволюция эмоций”. 2018
2. “Real-world adversarial attack on MTCNN face detection system” Edgar Kaziakhmedov, Klim Kireev, Grigorii Melnikov, Mikhail Pautov, Aleksandr Petiushko. Skolkovo Institute of Science and Technology; Moscow, Russia, arXiv:1910.06261v2 [cs.CV] 2 Apr 2020
3. “A Face Emotion Recognition Method Using Convolutional Neural Network and Image Edge Computing” Hongli Zhang, Alireza Jolfaei, And Mamoun Alazab, November 13, 2019. Digital Object Identifier 10.1109/ACCESS.2019.2949741
4. “Improved Kiwifruit Detection Using Pre-Trained VGG16 with RGB and NIR Information Fusion” Zhihao Liu, Jingzhu Wu, Longsheng Fu, Yaqoob Majeed, Yali Feng, Rui Li, And Yongjie Cui. IEEE Access, date of publication December 26, 2019. Digital Object Identifier 0.1109/ACCESS.2019.2962513
5. “Dropout: A Simple Way to Prevent Neural Networks from Overfitting” Nitish Srivastava Geoffrey Hinton Alex Krizhevsky Ilya Sutskever Ruslan Salakhutdinov. Department of Computer Science University of Toronto. Journal of Machine Learning Research 15 (2014) 1929-1958 Submitted 11/13; Published 6/14

V.P. Zelensky, V.E. Gai, P.A. Suvorkin

RECOGNIZING HUMAN EMOTIONS FROM A FACE IMAGE

Nizhny Novgorod state technical university n.a.R.E. Alekseev

The paper describes the process of creating an application for recognizing human emotions from a face image using deep learning algorithms.

Purpose: increasing the efficiency of face emotion recognition systems by creating modified convolution neural network.

Design/methodology/approach: machine learning algorithms, deep neural network.

Findings: the author has proposed new CNN modification and created face emotion recognition software.

Research limitations/implications: developed neural network can recognize only full face, but not in the profile.

Originality/value: This neural network can be used in client's feedback systems, or as base for other projects, connected with face photos and recognitions on them.

Keywords: face emotion recognition, machine learning, ML, FER, emotions, convolutional neural network, classification, VGG16, CNN, MTCNN, DNN

АППАРАТНО - ПРОГРАММНАЯ РОБОТОТЕХНИЧЕСКАЯ СИСТЕМА

Нижегородский государственный технический университет им. Р.Е. Алексеева

Целью работы является разработка аппаратно-программной робототехнической системы на примере системы управления прототипа беспилотного автомобиля, работа акцентируется на механизме объезда препятствий и ее калибровки.

Ключевые слова: ArduinoUno, ПИД-контроллер, рулевая рейка, ультразвуковые датчики, прототип беспилотного автомобиля

Аппаратно-программная робототехническая система может быть представлена в разнообразных видах и вариациях, для примера исследования данного аспекта технологии необходимо выбрать такую систему, которая послужит прототипом для актуальной технологии. Примером такой системы может служить система управления беспилотного автомобиля. При разработке данной системы требуется реализовать непрерывную, своевременную обработку входящих данных. Если программе понадобится слишком много времени, чтобы проанализировать показатели датчиков, рулевая рейка и передние колеса автомобиля будут двигаться с задержкой, то это приведет к потере контроля над автомобилем, например, при повороте руля в целях объезда препятствия машина не сразу среагирует на маневр. Подобная специфика повышает требования к временным показателям программ для автомобилей, особенно, если учесть ограниченную производительность процессора стандартных электронных блоков управления

Аппаратное обеспечение

Прототип беспилотного автомобиля можно разделить на несколько отдельных компонентов: Система датчиков, контроллер, обеспечение движения и маневров и питание.

Элементы питания состоит из двух частей: первая обеспечивает питание для электродвигателя, в виде трех аккумуляторов 18650 каждый на 4400 mAh, установлена под днищем автомобиля, вторая часть обеспечивает питание контроллеру через интерфейс USB в виде съемного аккумулятора на 5000mAh.



Рис. 1. Источник питания двигателя

Датчики в системе представлены двух типов:

1. ПИД-регулятор. Представляет собой датчик интенсивности света и круг на внутренней части передних колес со сменой цвета на черный и белый. Фиксируя

изменения цветов, вычисляется частота вращения колес и соответственно скорость движения автомобиля. На основе этих данных происходит управление скоростью во время выполнения маневров или при движении по прямой.

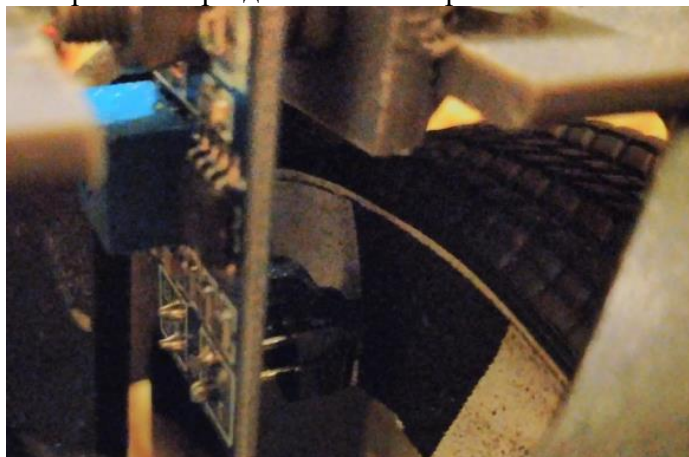


Рис. 2. ПИД-регулятор

2. Ультразвуковые датчики. Представлены в виде каскада датчиков на передней части автомобиля для обнаружения препятствий движению автомобиля. Датчики фиксируют препятствия на определенном расстоянии и подают сигнал на контроллер о необходимости начала выполнения маневра.

Контроллер состоит из двух микросхем: ArduinoUno, которая обрабатывает входящая данные от датчиков и передает соответствующие команды на систему движения и контроллер электродвигателя, который служит для обеспечения управления двигателем с помощью Arduino.

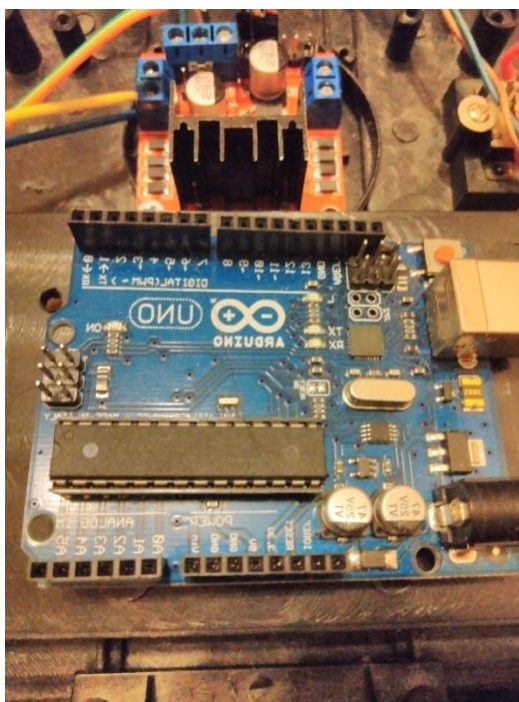


Рис. 3. Контроллер двигателя(сверху) и плата ArduinoUno(снизу)

Движение происходит благодаря электромотору, установленному на задней оси автомобиля, управление им происходит с помощью контроллера, описанному выше.

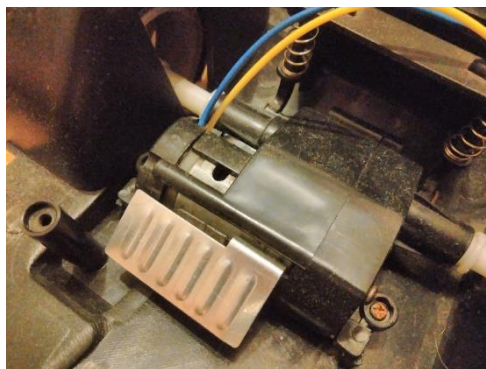


Рис. 4. Электромотор

Маневр обеспечивается с помощью рулевой рейки и сервопривода. При подаче сигнала о приближении к препятствию от ультразвуковых датчиков в зависимости от конфигурации программы подается команда на сервопривод и осуществляется поворот. Расположены данные элементы на передней оси автомобиля.



Рис. 5. Сервопривод

Программное обеспечение

Программная составляющая представлена из нескольких частей:

Обработка сигналов (с ПИД-регулятора и ультразвуковых датчиков) и управление движением на основе данных с датчиков (поворот и изменение скорости).

1. Изменение скорости (ПИД-контроллер): основываясь на показаниях ультразвуковых датчиков изменяется скорость движения, при обнаружении препятствия скорость снижается, в обратном случае соответственно остается неизменной или достигается стандартной (указанной в коде программы). Отслеживание скорости происходит через ПИД-регулятор.

2. Поворот осуществляется с помощью сервопривода, установленного на передней оси автомобиля, поворот будет осуществляться в сторону перемещения до цели и до тех пор, пока не будет препятствий движению машины.

3. Обработка сигналов:

ПИД-регулятор представляет собой датчик отслеживающий частоту смены черного и белого цветов на внутренней части передних колес. Соответственно мотор, установленный на заднюю ось будет увеличивать скорость пока не будет достигнута максимальная (установленная программой) скорость. При поступлении с ультразвуковых датчиков сигнала о приближении к препятствию скорость будет снижаться, до тех пока не будет достигнута установленная программой для подобной ситуации частота смены черного и белого на ПИД-регуляторе.

Ультразвуковые датчики. Располагаются на спереди по контуру габаритов машины, площадь покрытия представляет собой полукруг впереди автомобиля. При обнаружении препятствия будет изменяться скорость и осуществляться поворот до тех пор, пока препятствие будет вне поля зрения датчиков.

Эксперименты

Эксперименты необходимы для установления и проверки оптимальных параметров движения автомобиля: скорость по прямой, скорость при повороте, расстояние до препятствия при котором необходимо начать/закончить маневр. Для калибровки скорости движения будут проведены эксперименты следующего порядка:

1. Автомобиль разгоняется по прямой, перед ним на определенное расстояние устанавливается препятствие, которое необходимо объехать. Целью эксперимента является откалибровать максимальную скорость по прямой и расстояние, на котором должно фиксироваться препятствие для успешного объезда без столкновений. Зависимость этих двух параметров прямо пропорциональна: чем выше скорость движения до обнаружения преграды, тем на большем расстоянии она(преграда) должна быть обнаружена датчиками.

2. Устанавливается каскад препятствий, прототип должен пройти эти преграды за оптимальное время без предварительного разгона. В ходе эксперимента калибруется скорость маневра и расстояние, которое необходимо фиксировать датчикам. Аналогично предыдущему эксперименту была установлена прямо пропорциональная зависимость между этими параметрами: чем быстрее будет двигаться автомобиль при выполнении поворота, тем выше вероятность столкновения с препятствием, в следствие этого необходимо чтобы датчики фиксировали препятствия на большем расстоянии.

Заключение

Основные механизмы системы управления беспилотным автомобилем могут быть использованы во множестве разнообразных аспектов технологий, а при увеличении вычислительных мощностей становится возможным разработка более совершенных алгоритмов, например, использование технологий обработки изображений и компьютерного зрения может заменить ультразвуковые датчики тем самым расширив спектр задач, которые могут быть решены подобного рода технологий.

В данной статье были определены компоненты аппаратной и программной частей системы и требования к конечному результату.

Библиографический список

1. Уроки Arduino. ПИД регулятор // Youtube URL: <https://www.youtube.com/watch?v=rIbWnB26dp0>
2. Основы работы с Arduino // <https://microkontroller.ru/arduino-projects/kak-nachat-rabotu-s-arduino-uno-rukovodstvo-dlya-nachinayushhih/>
3. Основы разработки ПО для беспилотных автомобилей. Первая часть // <https://habr.com/ru/post/277141/>
4. Основы разработки ПО для беспилотных автомобилей. Вторая часть // <https://habr.com/ru/post/277315/>
5. Автомобильное программное обеспечение // https://press.ocenin.ru/avtomobilnoe-programmnoe-obespechenie/#Pazr_pro_na_baz
6. Описание плат ArduinoUno// https://press.ocenin.ru/avtomobilnoe-programmnoe-obespechenie/#Pazr_pro_na_baz

V.E. Gai, A.A. Tischenko

HARDWARE AND SOFTWARE OF THE ROBOTIC SYSTEM

Nizhny Novgorod State Technical University n.a. R.E. Alekseev

Purpose: Development of a hardware and software of the robotic system in the form of a prototype self-driving car, the work of system focuses on avoiding obstacles

Design/methodology/approach: open source

Findings: the author described the implementation of hardware and software of the robotic system, which includes the movement of a car and overcoming static obstacles

Originality/value: the driverless car control system is part of an extremely promising direction. The main mechanisms of this technology can be used in many aspects

Keywords: Arduino Uno, PID-controller, steering rack, Ultrasonic sensors, prototype of a self-driving car

ПРОГРАММНАЯ СИСТЕМА РАСПОЗНАВАНИЯ ПЕЧАТНОГО ТЕКСТА

Нижегородский государственный технический университет им. Р.Е. Алексеева

Целью работы является извлечение слов и строк из изображения, распознавание символов в словах и строках, формирование электронного документа или же блока текста для копирования

Ключевые слова: распознавание символов, детектирование слов

Первичным переводом документа в электронную форму является его сканирование или фотографирование, в результате которого получается графический файл в виде фотографии или скана. Однако такие файлы, особенно высокого разрешения, занимают много места на диске, и текст в них невозможно редактировать. В связи с этим, целесообразно извлекать текст из графических файлов, что успешно делается с применением OCR [1].

Оптическое распознавание символов (OCR) — перевод изображений машинописного, рукописного или печатного текста в электронные текстовые данные [2]. Системы OCR применяются во многих областях. Вот некоторые из задач, которые решают системы распознавания текстов [3]:

1. Считывание данных с бланков и анкет.
2. Автоматическое распознавание номерного знака.
3. Распознавание паспортных данных.
4. Извлечение информации из визитных карточек в список контактов.
5. Создание цифровых версий печатных и рукописных документов.
6. Технология для помощи слепым и слабовидящим.

С задачей распознавания символов связаны следующие проблемы:

1. Разнообразие форм начертания символов: документ может содержать несколько шрифтов сразу, а символы могут быть схожи по начертанию.
2. Искажение изображения, содержащего текст:
3. Шумы при печати.
4. Плохое качество изображения (засвеченность, размытость).
5. Вариации размеров, масштаба и положения символов на странице.
6. Влияние исходного масштаба печати: система оптического распознавания текста должна быть нечувствительной по отношению к способу верстки, расстоянию между строками и другим параметрам печати.

Процесс распознавания текста

Система распознавания текста предполагает наличие на входе изображения с текстом. На выходе система должна выдать текст, выделенный из входных данных [3, 4]. Весь процесс распознавания текста состоит из нескольких задач.

Обработка изображения

Перед началом распознавания текста изображение должно быть очищено от шума и приведено к виду, позволяющему эффективно выделять символы и распознавать их. Обычно у изображения повышают резкость, контрастность, выравнивают его и преобразовывают в используемый системой формат.

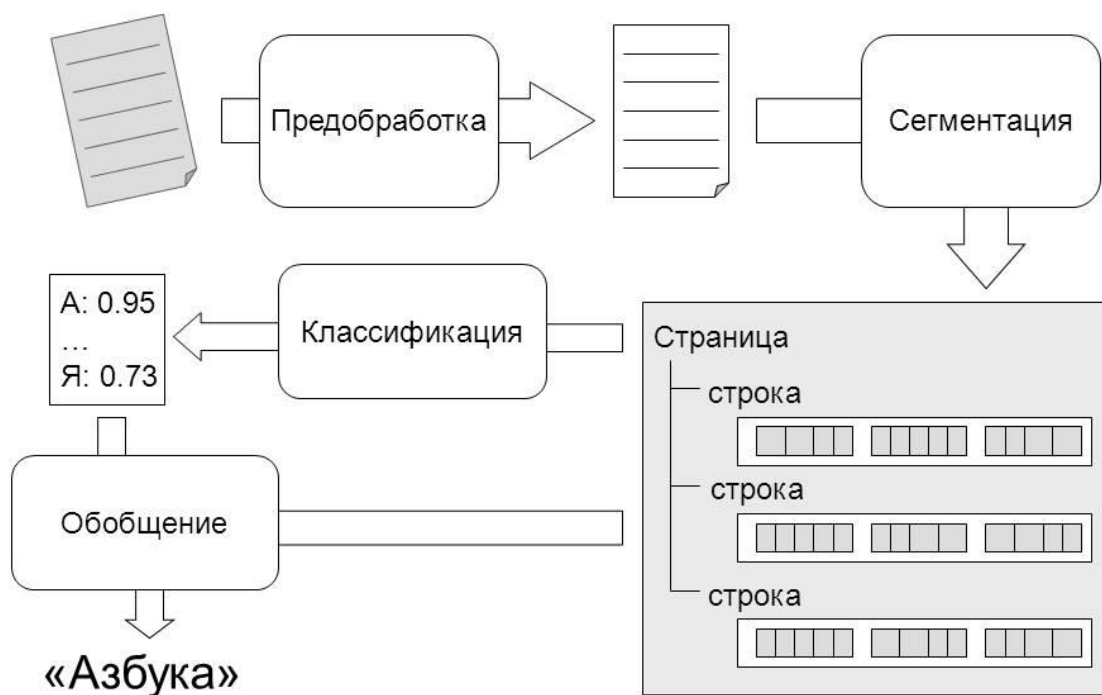


Рис. 1. Процесс распознавания текста

Алгоритмы распознавания символов

Существует несколько классов алгоритмов распознавания символов, но все они требуют высококачественного входного изображения [5]. Нечеткое изображение с большим количеством шумов будет обработано некорректно с большей вероятностью, чем высококонтрастное изображение без шума. Также на качество распознавания влияют наклон и размер символов. Поэтому перед использованием алгоритма распознавания необходимо предварительно обработать изображение. В настоящее время наиболее распространенными методами распознавания символов являются шаблонный метод, нейронные сети и некоторые другие методы.

Шаблонный метод

Одним из классических методов распознавания символов является шаблонный метод. В основе шаблонного метода лежит сравнение входного изображения и ряда распознанных символов шаблонных изображений. Размерность входного изображения и шаблонов должна быть одинаковой. Для этого необходимо произвести нормализацию входного изображения. Для идентификации символа вычисляются специальные коэффициенты корреляции, которые показывают степень различия между входным изображением и шаблоном. В этом случае, если один из коэффициентов корреляции намного важнее других, то это входное изображение четко идентифицируется.

Достоинства этого метода - простота реализации алгоритма, относительно высокая скорость распознавания, а также хорошее распознавание символов с различными дефектами, например, дроблеными или склеенными.

Недостаток этого метода в том, что для правильного распознавания необходимо писать символы по заранее заданному шаблону.

Ориентировочный метод

В ориентировочном методе создается N-мерный вектор признаков, например, количество закрытых областей, процент занятости и т. д. Построенный вектор признаков сравнивается с набором опорных векторов того же размера путем вычисления евклидова расстояния между опорным вектором и вектором признаков распознанного текста.

Преимущество этого метода - невосприимчивость к форме символа, его наклону, пропорциям и т. д.

Недостатком метода является неустойчивость к различным дефектам изображения.

Метод структурного распознавания

Метод структурного распознавания преобразует входное изображение в его топологическое представление, которое отображает информацию об относительном положении структурных элементов символа. Эта информация чаще всего представлена в виде графиков.

Этот метод имеет преимущества и недостатки, аналогичные ориентировочному методу.

Искусственная нейронная сеть

В последнее время одним из самых распространенных методов распознавания символов является распознавание с помощью нейронной сети. Основное отличие использования нейронной сети от классических методов распознавания состоит в том, что нейронные сети не программируются, а обучаются. Возможность обучения - одно из основных преимуществ нейронных сетей перед традиционными алгоритмами. Кроме того, этот метод отличается высокой эффективностью и производительностью. К недостаткам нейронных сетей можно отнести то, что данный метод распознавания требует большого объема данных, используемых для обучения, и разработки специальной структуры нейронной сети с учетом специфики этой задачи.

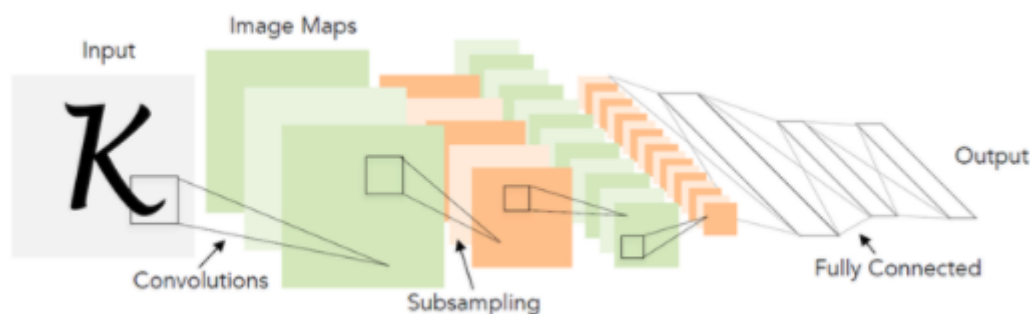


Рис. 2. Сверточная нейронная сеть для распознавания символа

Недостатки нейронных сетей

Нейронные сети с успехом могут применяться в системах распознавания текста, однако обладают существенными недостатками, препятствующими их широкому применению [3]:

1. Затраты памяти: необходимо построить достаточно большую сеть элементов, что приводит к большим затратам памяти.
2. Затраты ресурсов системы: в процессе распознавания используются большие объемы ресурсов системы, так как функции на элементах сети работают с числами с плавающей точкой.
3. Необходимость в обучении: для достижения более точного результата нейронную сеть необходимо обучать, однако и это не гарантирует идеальный результат.
4. Сложность построения: так как работа нейронной сети во многом зависит от ее конфигурации, требуется больше усилий для создания наиболее эффективной архитектуры.

Вычисления в сверточной сети

В сверточном слое входное исходное изображение или карта предыдущего сверточного блока подвергаются операции свертки при помощи матрицы небольшого размера, которую двигают по матрице, описывающей входные данные [1]. Выходными данными является матрица, состоящая из значений суммы попарных произведений соответствующей части входных данных с ядром свертки.

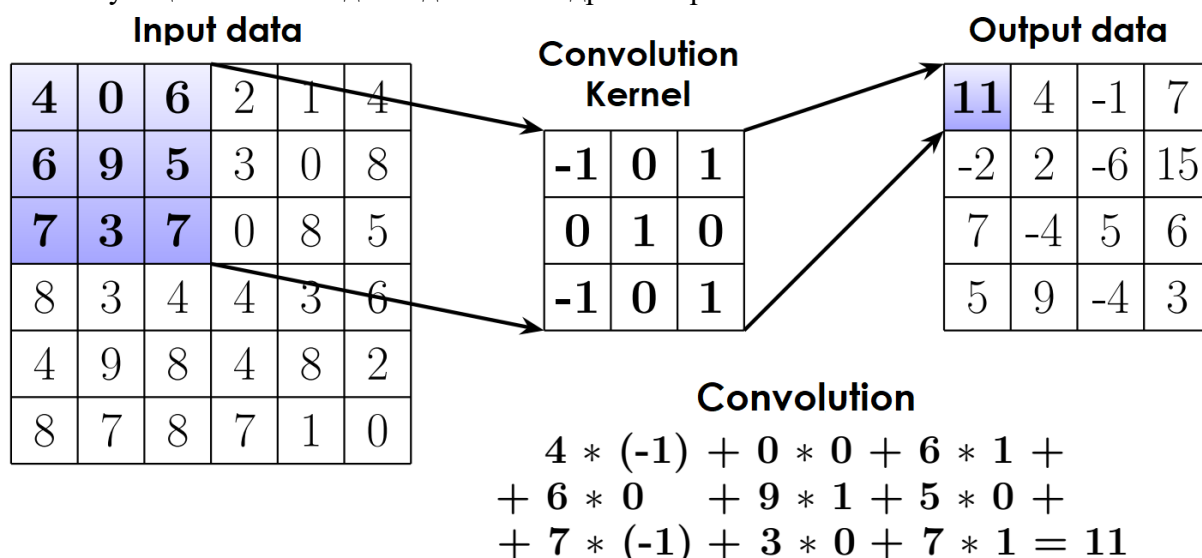


Рис. 3. Пример сверточного слоя

В слое пулинга уменьшается размерность выходных данных сверточного слоя в два этапа.

1. Применение функции активации нейронов. Наиболее используемой функцией активации в сверточных нейронных сетях является ReLU, которая заменяет отрицательные значения в матрице нулями. В последнее время доказали свою эффективность ее разновидности — NoisyReLU, которая заменяет отрицательные значения нулями с добавлением небольшого случайного слагаемого к положительным значениям, и LeakyReLU, которая заменяет отрицательные значения не нулями, а случайными числами.

2. Уменьшение размерности. Полученная после активации матрица разделяется на ячейки, в которых происходит агрегирование значений. Это можно сделать разными способами, но зачастую применяется метод выбора максимального элемента (max-pooling).

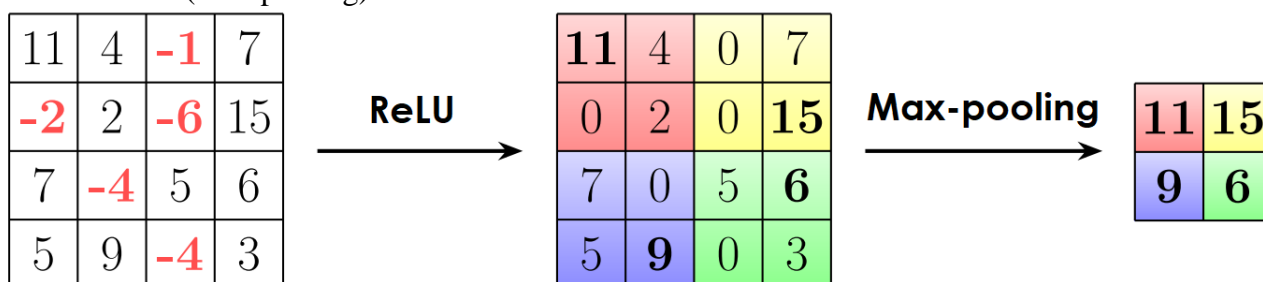


Рис. 4. Пример слоя пулинга

Заключение

Данная технология актуальна. Постоянно попадают документы, информация с которых необходима в электронном виде незамедлительно. Были рассмотрены алгоритмы распознавания символов, приведены все преимущества и недостатки.

Библиографический список

1. OCR-конвейер для обработки документов URL:
<https://habr.com/ru/company/arcadia/blog/505950/>
2. Оптическое распознавание символов URL:
https://ru.wikipedia.org/wiki/Оптическое_распознавание_символов
3. Распознавание текста на изображении URL:
http://neerc.ifmo.ru/wiki/index.php?title=Распознавание_текста_на_изображении
4. Ломанов Д. К. Алгоритм распознавания текстовой информации на изображении с помощью ЭВМ / Д. К. Ломанов. — Текст: непосредственный // Молодой ученый. — 2019. — № 28 (266). — С. 14-16.
5. Character recognition algorithms A A Valke and D G Lobov 2019 J. Phys.: Conf. Ser. 1210 012154

М.А. Pyin, V.E. Gai

PRINTED TEXT RECOGNITION SOFTWARE

Nizhny Novgorod state technical university n. a. R.E. Alekseeva

Purpose: extracting words and lines from an image, recognizing characters in words and lines, generating an electronic document or block of text for copying.

Design/methodology/approach: character recognition algorithms, minimal resources.

Findings: the author considered algorithms for character recognition, presented their advantages and disadvantages.

Research limitations/implications: recognition in only one language, minimal resources are used.

Originality/value: the app can be used on any available device.

Keywords: character recognition, word detection.

УДК 519.6

Галанов Н.Г.², Козелков А.С.¹⁻³, Жучков Р.Н.², Саразов А.В.²

ТЕСТИРОВАНИЕ МЕТОДИКИ МОДЕЛИРОВАНИЯ ПРОЦЕССА ОБЛЕДЕНЕНИЯ В ПАКЕТЕ ПРОГРАММ ЛОГОС

¹Нижегородский государственный технический университет им. Р.Е. Алексеева

²ФГУП «Российский Федеральный Ядерный Центр «Всероссийский научно-исследовательский институт экспериментальной физики»

³Саровский физико-технический институт – филиал федерального государственного автономного образовательного учреждения высшего образования «Национальный исследовательский ядерный университет «МИФИ»

В работе приводятся методы расчетов задач обледенения в модуле ЛОГОС-Аэро. Для верификации представлены результаты моделирования обледенения с использованием Эйлера и Лагранжевого подхода на задачах из верификационного базиса НАСА, предназначенного для тестирования программного пакета Lewice. С описанными подходами в модуле ЛОГОС-Аэро представляется возможность с высокой степенью точности моделировать различные формы льда, проводить моделирование процессов обледенения с определением формы ледяного нароста на различных летательных аппаратах.

Ключевые слова: многофазные течения, неструктурированные сетки, Эйлеров подход, Лагранжев подход, обледенение, уравнение Навье-Стокса.

Введение

Опыт эксплуатации авиационной техники показывает, что обледенение является одним из наиболее опасных воздействий естественной внешней среды и создает серьезную проблему для безопасности полетов гражданских самолетов во всем мире. Согласно статистическим данным более 50% летных происшествий происходит по причине обледенения элементов мотогондол двигателей, фюзеляжа или крыла [1]. При этом существенно ухудшаются аэродинамические характеристики крыла, увеличивается вес самолета и возрастает расход топлива. Наиболее опасными считаются режимы полета на небольших скоростях при прохождении самолета через зоны с большим содержанием переохлажденных капель воды. Формы, размеры наростов определяются при полетах в естественных условиях обледенения, при полетах за танкерами, разбрызгивающих воду, и при продувках элементов ЛА в специальных аэродинамических трубах, позволяющих моделировать обледенение [2]. Все приведенные испытания являются очень дорогостоящими, требуют много времени на их проведение и не позволяют провести исследование обледенения ЛА в широких пределах в короткие сроки. Поэтому для уменьшения трудозатрат на проведение экспериментов внедряются современные численные методики расчетов обледенения.

Степень исследованности

Данной тематикой занимаются большое количество исследовательских групп, как у нас в стране, так и за рубежом. Наиболее известными программными пакетами в мире, предназначенными для расчета задач обледенения, являются зарубежные ПО, такие как ANSYS FENSAP-ICE, LEWICE(2D), STAR-CCM+, ONERA. Связано это с тем, что данные программные пакеты являются основоположниками развития моделирования процессов обледенения. С начала 90-х годов число исследовательских групп, занимающихся

исследованием процесса обледенения, резко возросло. Исследованиями в области обледенения и влияния его на АДХ ЛА активно занимаются в России и СНГ во ФГУП «ЦАГИ», ФГУП «ЦИАМ», АО «ЛИИ», ООО «ТЕСИС». Представленными авторами разработаны различные программные коды, например, такие как: EWT ЦАГИ, IceVision, NSMB-ICE, NSCODE-ICE, авторский код Алексеенко С.В. и многие другие. В представленных программных кодах используются различные подходы и методы численного моделирования обледенения.

В настоящее время результаты трехмерных численных расчетов, полученные при решении уравнений Навье-Стокса на неструктурированных сеточных моделях, в достаточной степени согласуются с экспериментальными результатами. Принимая во внимание необходимость решения вышеописанных задач по развитию программных средств с использованием данного подхода в пакете программ ЛОГОС [3,5] были реализованы методы и алгоритмы моделирования процессов обледенения летательных аппаратов.

Подходы к моделированию частиц в расчетной области для процессов обледенения

Для моделирования движения капель воды в пакете программ ЛОГОС [3,5] применяются три подхода:

- первый с использованием Лагранжевой многофазности, в котором капли моделируются в дискретном приближении;
- второй с использованием Эйлеровой многофазности с представлением водяной фазы в приближении сплошной среды;
- третий подход является гибридным - моделирование жидкой фазы проводится Эйлеровским подходом, а образование и растекание пленок на твердых поверхностях на основе Лагранжевого подхода.

В Лагранжевом подходе математическое моделирование двухфазного потока основано на численном решении основных уравнений сохранения для газовой и конденсированной фаз. При этом конденсированная фаза рассматривается с точки зрения метода квазичастиц. Этот метод подразумевает решение дифференциальных уравнений для траектории, количества движения, энергии и массы ансамбля квазичастиц, которые представляют собой группу невзаимодействующих между собой индивидуальных частиц (капель) с одинаковыми физическими свойствами. Использование Лагранжевой многофазности для моделирования задач обтекания объектов большой размерности до сих пор является сложной задачей.

В Эйлеровом подходе исходят из предположения, что газовый поток и поток капель связаны в одностороннем порядке (жидкая фаза не влияет на поток газа), при этом все капли равномерно распределены в расчетной области. Данная модель не рассматривает такие процессы как испарение и конденсация капель. Использование Эйлеровой многофазности менее трудоемко.

Методики расчета задач обледенения

Задачи моделирования обледенения в ПК ЛОГОС проводятся с использованием трех типов расчетов:

- одностадийный;
- многостадийный;
- нестационарный.

Одностадийный тип расчета предполагает получение стационарных характеристик газодинамических параметров и новой формы поверхностей, подверженных обледенению, с помощью процедуры морфинга расчетной сетки.

Методика многостадийного вычисления подразделяется на отдельные стадии, которые содержат следующую последовательность решения задач:

- получение стационарного решения газового течения в расчетной области, включая расчет касательных напряжений на границе с твердыми стенками и тепловых потоков между газом и стенками;
- расчет стационарного движения капель в приближении Эйлера на основе полученного газодинамического решения;
- расчет толщины жидкой пленки и льда в течение заданного времени;
- получение новой формы поверхностей, подверженных обледенению, и изменение расчетной сетки.

На каждой последующей стадии используется расчетная сетка, измененная из-за образования льда на предыдущей стадии. Окончание расчета определяется заданным общим временем обледенения.

В случае нестационарного подхода проводится расчет газодинамического течения, движения капель и обледенения с единым шагом по времени без получения стационарных решений. При этом возможна передача источников массы, импульса и энергии от капель и жидких пленок в газодинамический решатель, что обеспечивает связь решателей между собой. На рис. 1 представлена блок-схема проведения расчетов в ПК ЛОГОС.

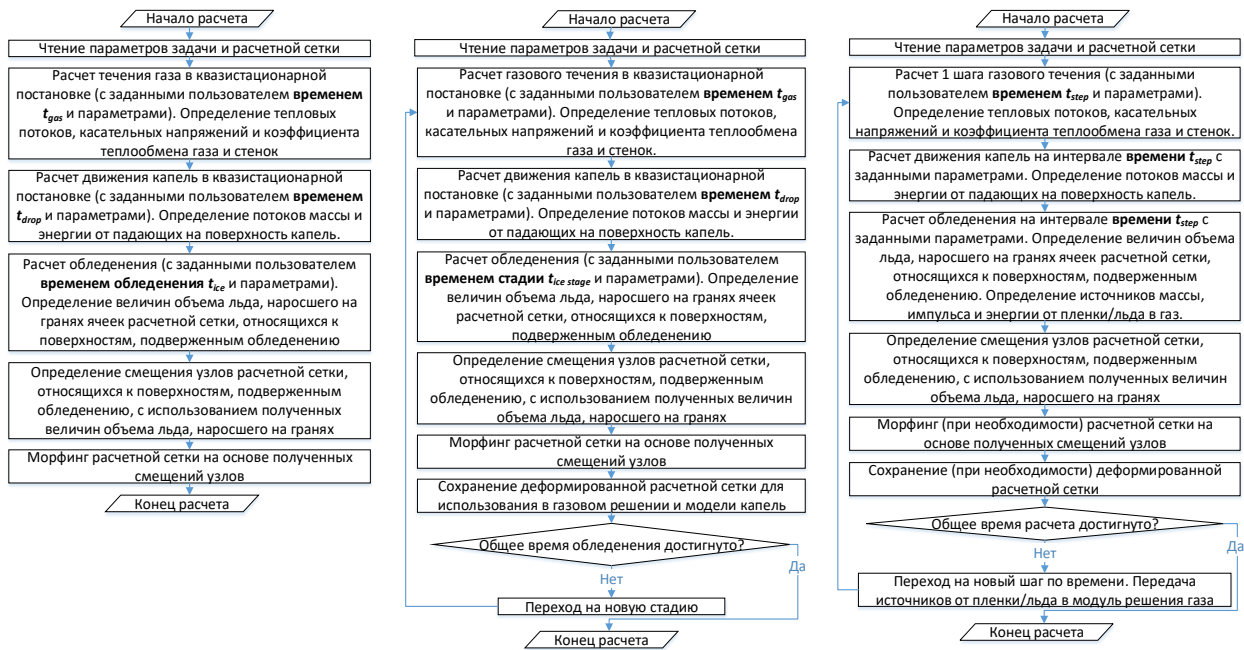


Рис. 1. Блок схемы проведения расчетов в ПК ЛОГОС

Верификация

Тестирование и верификация реализованных методик расчета задач обледенения в ПК ЛОГОС [5] проводилось на различных задачах. В данной работе приводятся результаты моделирования задачи обледенения профиля крыла NASA23014m, из верификационного базиса Lewice [4] и сравнение с результатами эксперимента. На следующем рис. 2 представлена расчетная сетка близи профиля крыла NASA23014m используемая для тестирования.

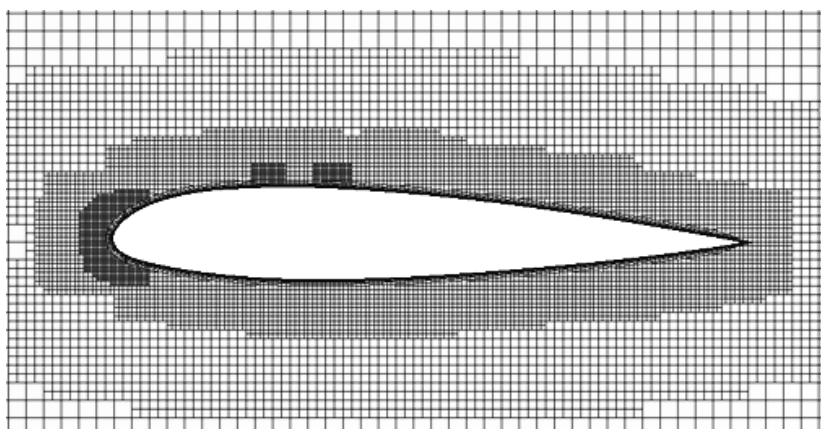


Рис. 2. Расчётная геометрия вблизи профиля

Расчётная сетка является структурированной и состоит ~19тыс ячеек.

Расчёт обледенения данного профиля крыла проводился на следующих режимах, представленных в табл. 1.

Таблица 1

Режимы моделирования процесса обледенения

№	Название профиля крыла	Хорда, м	№ эксп.	V , м/с	T_{tot}, K	T, K	$\alpha, ^\circ$	LWC , кг/м ³	MVD , м	Время, мин
1	NACA23013(mod)	1,745	219	87,2	255,37	251,3	0	8,2e-4	160e-6	3
2	NACA23013(mod)	1,745	251	87,2	270,93	266,85	0	8,2e-4	160e-6	21,2
3	NACA23013(mod)	1,745	123r8	88,5	272,04	267,85	5	8e-4	20e-6	10

Здесь V – скорость полёта;

T_{tot} – полная температура воздуха;

T – «Outside Air Temperature», статическая температура воздуха;

α – угол атаки;

LWC – «Liquid water content», водность потока;

MVD – «Median Volumetric diameter», средний объёмный диаметр капель;

Время – продолжительность процесса обледенения.

На следующих рисунках 3 - 5 представлены результаты моделирования обледенения на различных режимах, представленных в табл. 1.

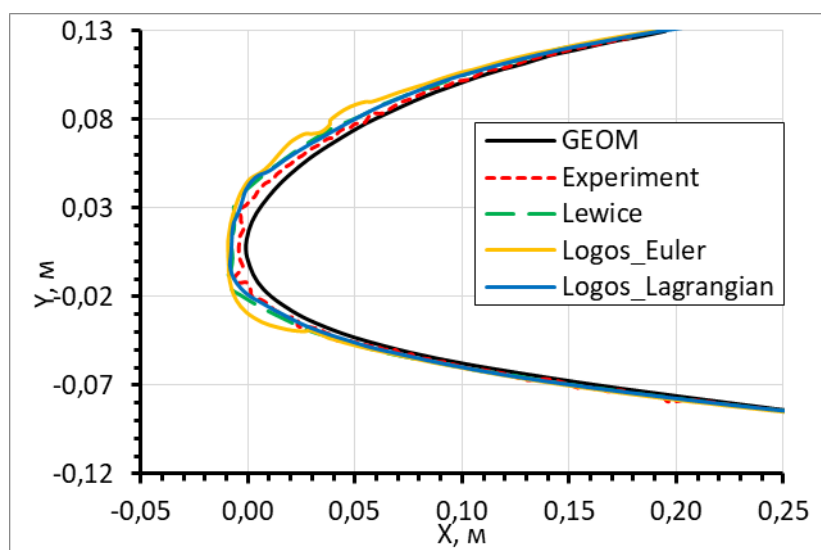


Рис. 3. Сравнение результатов обледенения для первой постановки

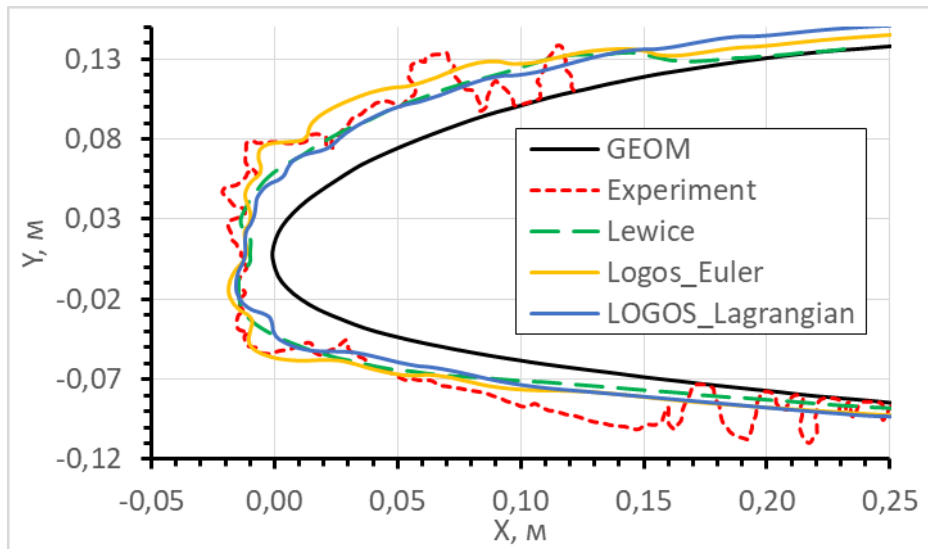


Рис. 4. Сравнение результатов обледенения для второй постановки

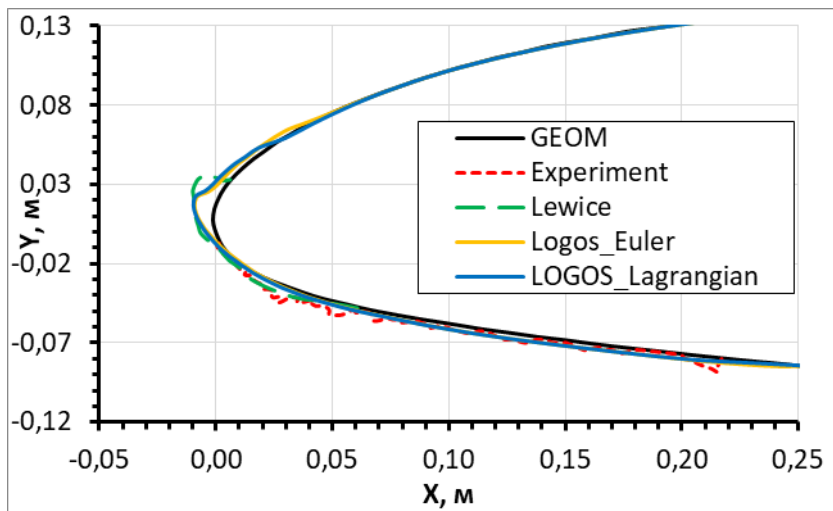


Рис. 5. Сравнение результатов обледенения для третьей постановки

На следующих рис. 6, 7 представлены результаты моделирования обледенения при многостадийном и многоэтапном режимах с использованием Эйлеравого и Лагранжевого подходов.

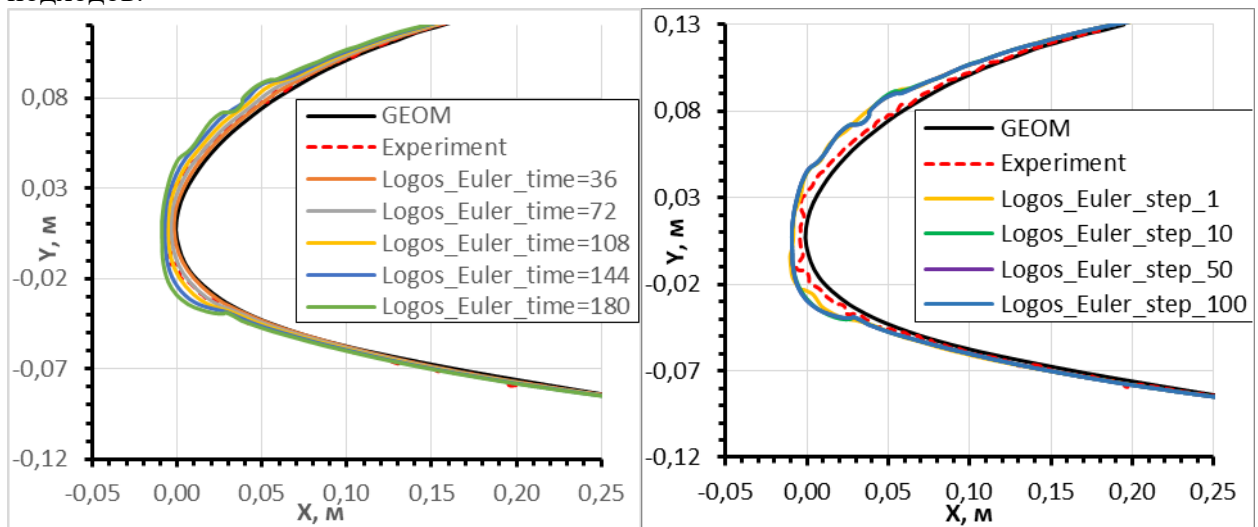


Рис. 6. Нарастание льда на профиль крыла NASA23014 (mod) при многостадийном подходе (слева) и при различном количестве этапах (справа) в Эйлеравом подходе

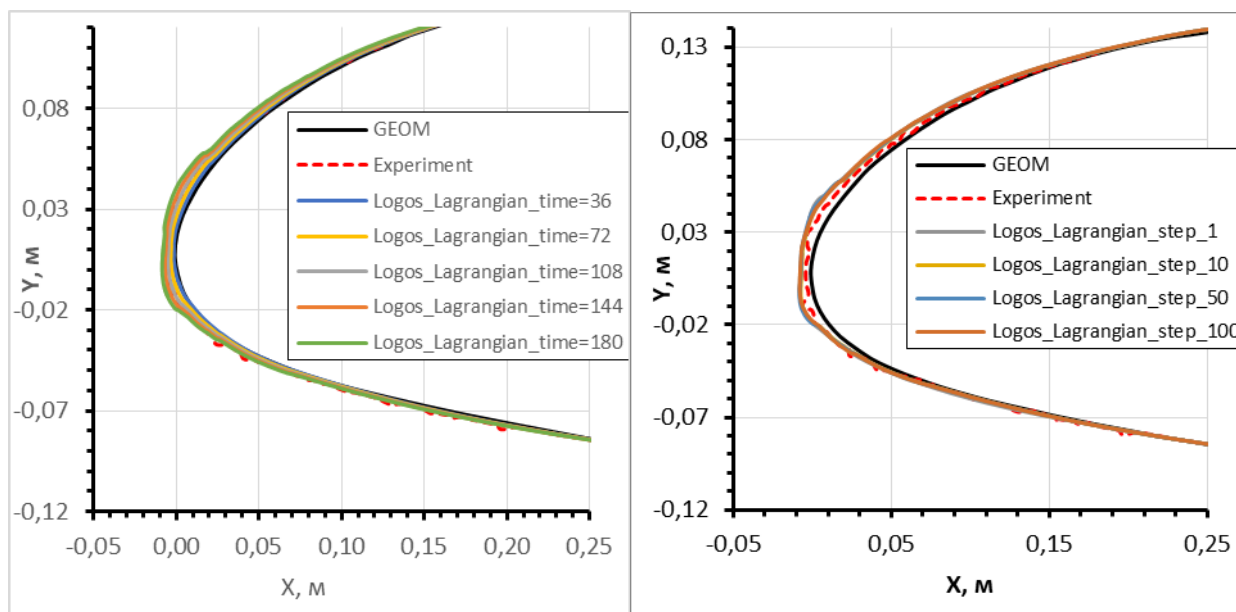


Рис. 7. Нарастание льда на профиль крыла NASA23014 (mod) при многостадийном подходе (слева) и при различном количестве этапах (справа) в Лагранжевом подходе

Как видно из представленных графиков форма ледяного нароста согласуется с экспериментальными данными и данными ПК Lewice [4].

Заключение

В модуль сжимаемых течений пакета программ ЛОГОС внедрены алгоритмы и модели, используемые для моделирования процесса обледенения ЛА. Реализованная методика позволяет проводить моделирование процессов обледенения с определением формы ледяного нароста на различных летательных аппаратах. В работе приводятся блок-схема проведения расчетов в ПК ЛОГОС входящие в данную методику. На некоторых задачах из верификационного базиса NASA, предназначенного для тестирования программного пакета Lewice [4], демонстрируется работоспособность реализованных компонент пакета программ ЛОГОС с использованием Лагранжевой и Эйлеровой многофазности и позволяющих с моделировать различные формы льда.

Представленные результаты получены при поддержке грантов Президента РФ по государственной поддержке ведущих научных школ РФ НШ-2485.2020.5.

Библиографический список

1. Ice accretion simulation // AGARD-AR-344. – 1997. – p. 280.
2. Сорокин К.Э., Бывальцев П.М., Аксенов А.А., Жлуктов С.В., Савицкий Д.В., Бабулин А.А., Шевяков В.И., Численное моделирование обледенения в программном комплексе FlowVision // Компьютерные исследования и моделирование. – 2020. – Т. 12, №1. – С. 83 – 96.
3. Козелков А.С., Куркин А.А., Шарипова И.Л., Курулин В.В., Пелиновский Е.Н., Тятюшкина Е.С., Мелешкина Д.П., Лашкин С.В., Тарасова Н.В. Минимальный базис задач валидации методов расчета течений со свободной поверхностью // Труды НГТУ им. Р.Е. Алексеева. – 2015. – № 2 (109). – С. 49-69.
4. Wright B.W., Rutkowski A. «Validation Results for LEWICE 2.0», NASA/CR--1999-208690, 1999.
5. Козелков А.С., Мелешкина Д.П., Куркин А.А., Тарасова Н.В., Лашкин С.В., Курулин В.В. Полностью неявный метод решения уравнений Навье-Стокса для расчета многофазных течений со свободной поверхностью // Вычислительные технологии. – 2016. – Т. 21, №5. – С. 54-76.

N.G. Galanov², A.S. Kozelkov¹⁻³, R.N. Zhuchkov², A.V. Sarazov²

**TESTING OF AN ICE ACCRETION SIMULATION TECHNIQUE IMPLEMENTED IN
THE SOFTWARE PACKAGE LOGOS**

¹*Nizhny Novgorod State Technical University n. a. R. E. Alekseev*

²*FSUE "Russian Federal Nuclear Center – All-Russian Research Institute
of Experimental Physics"*

³*SarPhTI MEPHI*

The paper presents ice accretion simulation methods implemented in the LOGOS-Aero module. For verification purposes, results of Eulerian and Lagrangian ice accretion simulations are reported for NACA verification problems, which are used to test the software package Lewice. The approaches described in the paper allow LOGOS-Aero to provide high-accuracy simulations of various ice shapes and ice accretion processes and predict the shape of ice deposits on various flight vehicles.

Keywords: multi-phase flow, unstructured grids, Eulerian approach, Lagrangian approach, ice accretion, Navier-Stokes equation.

ОДЕЛИРОВАНИЕ ПЛАВАНИЯ ТЕЛ НА БАЗЕ МЕТОДА VOF СОВМЕСТНО С МЕТОДОМ ДЕФОРМАЦИИ СЕТКИ

¹Федеральное государственное унитарное предприятие «Российский Федеральный Ядерный Центр Всероссийский научно-исследовательский институт экспериментальной физики» (ФГУП «РФЯЦ - ВНИИЭФ»), г. Саров

²Нижегородский государственный технический университет им. Р. Е. Алексеева, г. Н. Новгород

Рассматривается методика моделирования задач плавания тел, основанная на решении системы уравнений Навье-Стокса совместно с методом VOF. Учет подвижных тел осуществляется путем деформации расчетной сетки. Методика реализована на базе отечественного пакета программ ЛОГОС. Исследуется влияние численного метода и сеточных параметров на точность решения на основе задачи свободных затухающих колебаний цилиндра на водной поверхности. Результаты, полученные с помощью настоящей методики, сравниваются с аналитическими решениями и экспериментальными данными. Показано, что предлагаемая численная методика позволяет получать качественные результаты.

Ключевые слова: вычислительная гидродинамика, volume of fluid, плавание тел, свободная поверхность, уравнения Навье-Стокса.

Процесс плавания тел встречается во многих отраслях науки и техники, в первую очередь, в задачах судостроительной промышленности. Главные особенности задач данного класса – наличие свободной поверхности и подвижных твердых тел.

Рассматривается математическая модель, основанная на системе уравнений Навье-Стокса совместно с методикой VOF [1] и реализованная на базе отечественного пакета программ ЛОГОС [2-3]. Предполагается, что течение изотермическое, а поле скорости общее для всех фаз. С учетом этих допущений получается система уравнений, состоящая из уравнения сохранения массы, уравнения сохранения импульса и уравнения переноса объемной доли, которая в декартовых координатах имеет вид:

$$\begin{cases} \frac{\partial u_i}{\partial x_i} = 0, \\ \frac{\partial u_i}{\partial t} + \rho \frac{\partial}{\partial x_j} (u_i u_j) = -\frac{\partial p}{\partial x_i} + \frac{\partial}{\partial x_j} \tau_{ij} + \rho g_i, \\ \frac{\partial \alpha_w}{\partial t} + \frac{\partial}{\partial x_i} (u_i \alpha_w) = 0, \end{cases} \quad (1)$$

где i, j – нижние индексы, указывающие на принадлежность векторных компонент к декартовым координатам, $i, j = \{x, y, z\}$, ρ – средняя плотность, вычисляемая как $\rho = (\rho_w \alpha_w + \rho_a \alpha_a)$, w – (water) нижний индекс, указывающий на принадлежность к фазе «вода», a – (air) нижний индекс, указывающий на принадлежность к фазе «воздух», α_w – объемная доля воды, u_i – компонента вектора скорости, $i = \{x, y, z\}$, t – время, p – давление, x_i – компонента вектора декартовых координат, $i = \{x, y, z\}$, τ_{ij} – тензор вязких напряжений, который, согласно гипотезе Буссинеска, принимает вид:

$$\tau_{ij} = \mu \left(\frac{\partial u_i}{\partial x_j} + \frac{\partial u_j}{\partial x_i} + \frac{2}{3} \frac{\partial u_k}{\partial x_k} \delta_{ij} \right), \quad (2)$$

где μ – динамическая вязкость, δ_{ij} – символ Кронекера, g_i – компонента вектора ускорения свободного падения.

Учет силы тяжести производится с использованием алгоритма, основанного на поправке объёмных сил [4], который обеспечивает отсутствие паразитных осцилляций, связанных с неколокированным размещением неизвестных величин, на сетках любого типа.

Для учёта подвижных тел используется метод деформации расчётной сетки без изменения её топологии. Деформация сетки осуществляется методом IDW [5]. Учёт движения сетки в системе (1) осуществляется путём перехода в уравнениях переноса объёмной доли фаз и импульса к подвижной системе координат по известному закону [6]:

$$\frac{\partial^* \varphi}{\partial t} = \frac{\partial \varphi}{\partial t} + v_i \frac{\partial \varphi}{\partial x_i}, \quad (3)$$

где $\frac{\partial^* \varphi}{\partial t}$ – субстанциональная производная переносимого скаляра φ относительно подвижной системы координат, v_i – вектор скорости перемещения сетки. С использованием (3) уравнение переноса объёмной доли можно записать следующим образом:

$$\frac{d^* \alpha_w}{dt} + (u_i - v_i) \frac{\partial \alpha_w}{\partial x_i} + \alpha_w \frac{\partial u_i}{\partial x_i} = 0. \quad (4)$$

Здесь $\frac{d^* \alpha_w}{dt}$ – субстанциональная производная на движущейся сетке.

Уравнение сохранения импульса также формулируется относительно подвижной системы координат с учётом (3):

$$\rho \frac{d^* \alpha u_i}{dt} + \frac{\partial}{\partial x_j} (\rho u_i (u_j - v_j)) - u_i \frac{\partial}{\partial x_i} (\rho u_i (u_j - v_j)) = -\frac{\partial p}{\partial x_i} + \frac{\partial \tau_{ij}}{\partial x_j} + \rho g_i. \quad (5)$$

Уравнение сохранения массы записывается относительно скорости в подвижной системе координат:

$$\frac{\partial (u_i - v_i)}{\partial x_i} = 0. \quad (6)$$

Таким образом, конечная система уравнений (1) принимает следующий вид:

$$\begin{cases} \frac{\partial (u_i - v_i)}{\partial x_i} = 0, \\ \rho \frac{d^* \alpha u_i}{dt} + \frac{\partial}{\partial x_j} (\rho u_i (u_j - v_j)) - u_i \frac{\partial}{\partial x_i} (\rho u_i (u_j - v_j)) = -\frac{\partial p}{\partial x_i} + \frac{\partial \tau_{ij}}{\partial x_j} + \rho g_i, \\ \frac{d^* \alpha_w}{dt} + (u_i - v_i) \frac{\partial \alpha_w}{\partial x_i} + \alpha_w \frac{\partial u_i}{\partial x_i} = 0. \end{cases} \quad (7)$$

В таком виде уравнения системы (7) легко реализуемы в рамках конечно-объёмной технологии дискретизации [7-9].

В рамках данной работы исследуется влияние сеточного разрешения, величины шага по времени, порядка аппроксимации по времени и по пространству на точность моделирования плавания тел. В качестве базовой рассматривается задача о затухающих свободных колебаниях цилиндра на водной поверхности [10, 11].

Рассматриваемая задача представляет собой двумерный тест, расчетная область состоит из бассейна длиной $130d$ и высотой $16d$, заполненного водой на $8.17d$, и цилиндра диаметром $d = 0.1524$ м, плотность которого равна половине плотности воды. В начальный момент времени цилиндр погружен в воду на $0.167d$ выше своей оси, затем под действием силы Архимеда он начинает совершать свободные затухающие колебания на водной поверхности. Затухание движения обусловлено вязкостью и эффектом поверхностного натяжения. Для учета сил поверхностного натяжения используется модель CSF (continuum surface force) [12].

Сравнение результатов численного моделирования с экспериментальными [10] и аналитическими данными [11] по перемещению центра масс цилиндра от времени

приведено на рис.1. Результаты, полученные с помощью ПП ЛОГОС, хорошо согласуются с экспериментом и аналитическим решением.

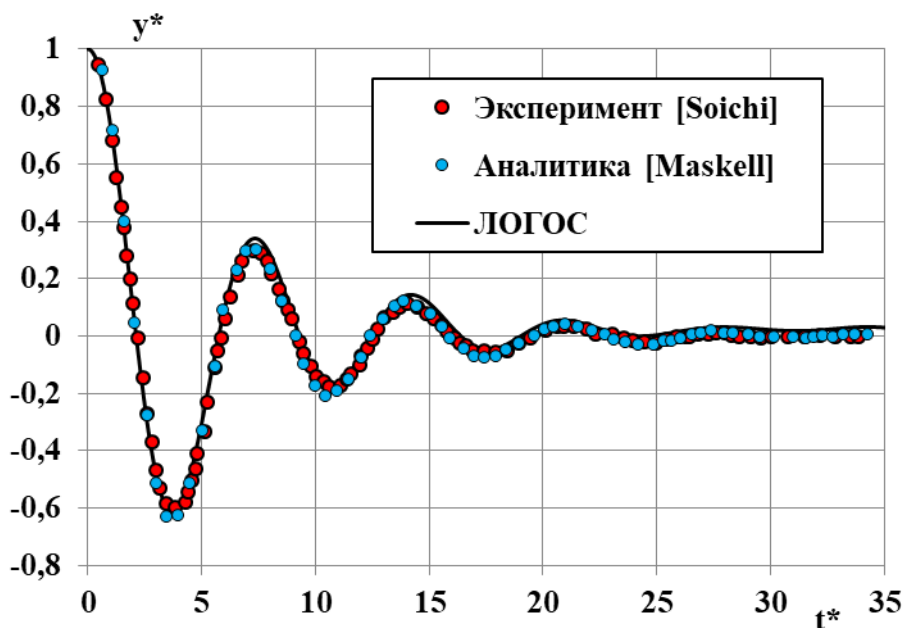


Рис. 1. Зависимость положения центра тяжести цилиндра от времени

Для исследования влияния сеточного разрешения на точность получаемого результата расчет проводится на четырех сетках, отличающихся базовым размером ячейки и толщиной призматического слоя. В качестве сравнительной характеристики рассматривается перемещение центра масс во времени. Используются нормированные координаты $y^* = y/y_0$ и время $t^* = t\sqrt{g/r}$, где y_0 – начальная координата центра масс цилиндра, g – ускорение свободного падения, r – радиус цилиндра. Результаты расчета на четырех различных сетках показаны на рисунке 2.

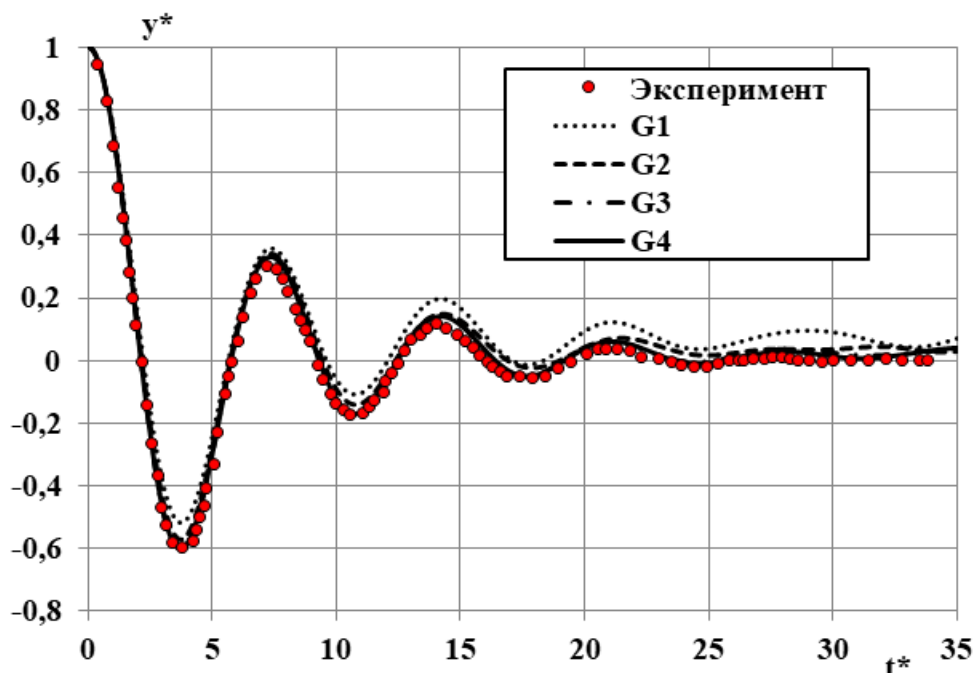


Рис. 2. График зависимости координаты цилиндра от времени для G1-G4 и схемы UD

В качестве количественной меры выбираются значения двух экстремумов временной зависимости перемещения центра масс тела, которые соответствуют всплытию цилиндра – M1, M2. Отклонения значений экстремумов, полученных в расчетах со схемой UD и LUD на различных сетках, от экспериментального значения приведены в табл. 1.

Таблица 1

Отклонения значений экстремумов, полученных на сетках G1-G4, от эксперимента

Экстремум	Отклонение от эксперимента, %							
	G1		G2		G3		G4	
	UD	LUD	UD	LUD	UD	LUD	UD	LUD
M1	13.4	11.4	4.4	3.1	-1.3	-2.8	-0.5	-1.0
M2	38.6	33.4	20.2	19.2	3.6	1.1	3.2	1.6

Исследование показало, что при различном сеточном разрешении максимальное отклонение решения от референсных данных изменяется пропорционально изменению базового размера ячейки расчетной сетки. Использование схемы второго порядка дискретизации конвективных слагаемых для скорости позволяет увеличить точность получаемого решения.

Для исследования влияния на результат величины временного шага и схемы дискретизации по времени рассматривается шесть различных значений временных шагов. На рисунке 3 представлены зависимости положения центра тяжести цилиндра от времени для расчетов с различными временными шагами и схемой дискретизации по времени второго порядка.

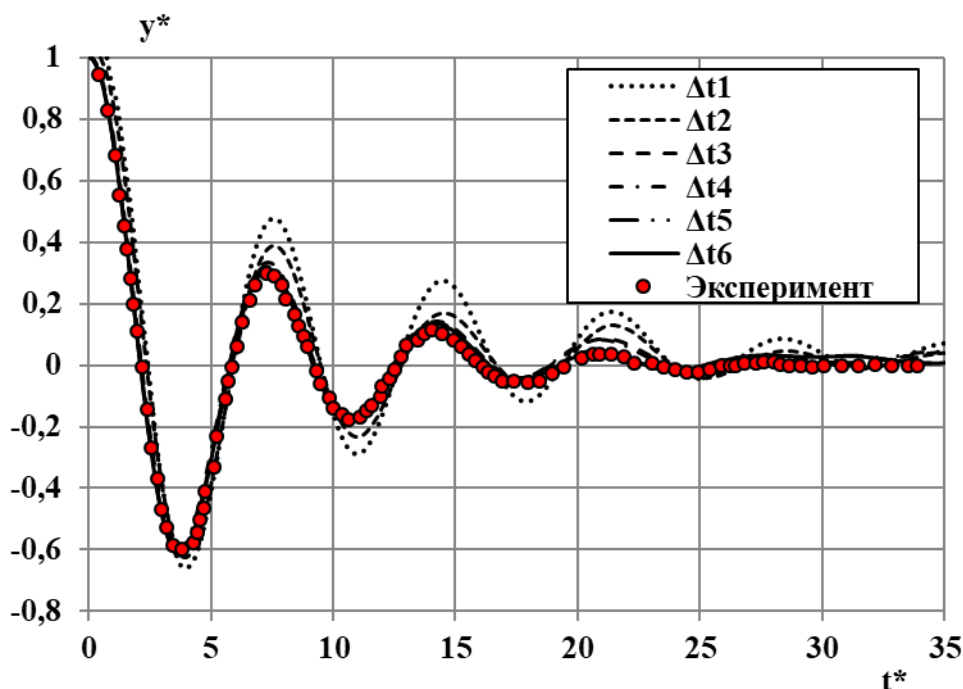


Рис. 3. Сравнение перемещения центра масс цилиндра при различных временных шагах

Отклонение значений экстремумов, полученных в расчетах с различными временными шагами и порядком по времени, от экспериментального значения приведено в табл. 2.

Отклонение значений экстремумов, в расчетах с различными временными шагами и порядком по времени, от эксперимента

Экстремум	Отклонение от эксперимента, %											
	Δt_1^*		Δt_2^*		Δt_3^*		Δt_4^*		Δt_5^*		Δt_6^*	
	I	II	I	II	I	II	I	II	I	II	I	II
M1	-4.4	-10.5	-4.5	-4.6	1.0	-0.4	0.2	-0.1	-1.6	-1.6	-1.6	-1.7
M2	-9.2	-64.1	-6.5	-32.5	0.2	-3.3	-0.3	-2.9	-3.8	-3.6	-4.5	-3.6

Сравнение результатов расчетов с различными временными шагами показало, что при большом шаге схема дискретизации по времени первого порядка дает меньшее отклонение от референсных данных, однако при уменьшении шага по времени до определенного достаточно малого значения использование схемы второго порядка позволяет получить более точный результат.

Представленные результаты получены при поддержке грантов Президента РФ по государственной поддержке ведущих научных школ РФ НШ-2485.2020.5.

Библиографический список

1. Hirt C.W., Nichols B.D. Volume of fluid (VOF) method for the dynamics of free boundaries // J. Comput. Phys., 1981, Vol. 39, p. 201-225.
2. Deryugin Yu., Zhuchkov R., Zelenskiy D., Kozelkov A., Sarazov A., Kudimov N., Lipnickiy Yu., Panasenko A. Validation results for the LOGOS multifunction software package in solving problems of aerodynamics and gas dynamics for the lift-off and injection of launch vehicles // Mathematical Models and Computer Simulations, 2015, V.7. Issue 2, P. 144–153.
3. Курулин В.В., Козелков А.С., Ефремов В.Р., Яцевич С.В., Тарасова Н.В. Применение метода VOF для решения комплексных задач со свободной поверхностью // XXV Всероссийский семинар с международным участием по струйным, отрывным и нестационарным течениям, 2018, С. 144-145.
4. Ефремов В.Р., Козелков А.С., Корнев А.В., Куркин А.А., Курулин В.В., Стрелец Д.Ю., Тарасова Н.В. Метод учета сил гравитации при моделировании течений со свободной поверхностью // Журнал вычислительной математики и математической физики, 2017, Т. 57, № 10, С. 1748-1762.
5. Luke Edward, Collins Eric, Blades Eric A fast mesh deformation method using explicit interpolation // Journal of Computational Physics, 2012, № 231, P. 586–601.
6. Лойцянский Л.Г. Механика жидкости и газа. – М: Государственное издательство технико-теоретической литературы, 1950.
7. Козелков А.С., Курулин В.В., Лашкин С.В., Шагалиев Р.М., Ялозо А.В. Исследование потенциала суперкомпьютеров для масштабируемого численного моделирования задач гидродинамики в промышленных приложениях // Журнал вычислительной математики и математической физики, 2016, Т. 56, № 8, С. 1524–1535.
8. Флетчер К. Вычислительные методы в динамике жидкостей в двух томах. – М: Мир, 1991.
9. Jasak H. Error Analysis and Estimation for the finite volume method with applications to fluid flows. Thesis submitted for the degree of doctor // Department of Mechanical Engineering, Imperial College of Science, 1996.
10. Soichi Ito, Study of the transient heave oscillation of a floating cylinder // Massachusetts institute of technology, 1977.
11. Maskell, S. J., and Ursell, F. The transient motion of a floating body // J. Fluid Mech., 1970, V. 44, P. 303-313.
12. Brackbill J.U., Kothe D.B., Zemach C. A continuum method for modeling surface tension // J. Comput. Phys., 1992, V. 100, P. 335-354.

Plygunova K.S.¹, Kozelkov A.S.^{1,2}, Kurulin V.V.¹, Utkin D.A.¹

**SIMULATIONS OF FLOATING BODIES BASED ON A COMBINATION OF THE VOF
AND MESH MORPHING METHODS**

*¹Federal State Unitary Enterprise “Russian Federal Nuclear Center All-Russian
Research Institute of Experimental Physics”*

²Nizhny Novgorod State Technical University n. a. R. E. Alekseev

The paper considers a floating-body simulation technique based on a combination of the Navier-Stokes equations and the VOF method. Moving bodies are modeled by mesh morphing. The simulation technique has been implemented within the domestic software package Logos. Sensitivity of solution accuracy to the numerical method and mesh parameters is studied for the case of free decaying oscillations of a cylinder floating on the water surface. Results calculated by our technique are compared with analytical solutions and experimental data. We show that our technique can produce high-quality results.

Keywords: computational fluid dynamics, volume of fluid, body floating, free surface, Navier-Stokes equations

ЧИСЛЕННОЕ ТРЕХМЕРНОЕ МОДЕЛИРОВАНИЕ ЗАДАЧ ГИДРОДИНАМИКИ НА СЕТКАХ, СОДЕРЖАЩИХ НЕСОГЛАСОВАННЫЕ СЕТОЧНЫЕ ИНТЕРФЕЙСЫ

¹*Федеральное государственное унитарное предприятие «Российский Федеральный Ядерный Центр Всероссийский научно-исследовательский институт экспериментальной физики» (ФГУП «РФЯЦ - ВНИИЭФ»), Нижегородская обл., г. Саров*

²*Нижегородский государственный технический университет им. Р. Е. Алексеева, г. Нижний Новгород*

В данной работе представлено описание численного метода, учитывающего особенности решения уравнений гидродинамики при моделировании течений вязкой несжимаемой жидкости в области смежных границ несогласованных фрагментов сетки. Описан пример реализации метода для случая аппроксимации уравнений Навье-Стокса методом конечных объемов. В основу метода положен принцип работы GGI (General Grid Interface) не требующий модификации исходной сетки и основанный на консервативной интерполяции потоков в области несогласованных интерфейсов. Представлены результаты решения задачи турбулентного течения в круглой трубе с внезапным сужением.

Ключевые слова: General Grid Interface, ЛОГОС, несогласованный сеточный интерфейс.

Введение

В настоящее время значительно возросла потребность в численном моделировании гидродинамики, основанном на решении трехмерных уравнений Навье-Стокса произвольных неструктурированных сетках большой размерности [1]. В большинстве случаев для решения таких задач строится единая сетка, однако разбиение сложной исходной геометрии на более простые фрагменты значительно упрощает задачу построения, позволяет получить сетки лучшего качества и значительно сэкономить вычислительные ресурсы. Как правило построенные таким образом сеточные модели состоят из несогласованных фрагментов сетки, на смежных границах которых количество и расположение узлов в общем случае не совпадает.

Ключевой особенностью численного моделирования задач гидродинамики на таких сеточных моделях является необходимость связывания несогласованных фрагментов сетки в единую расчетную модель с помощью специальных несогласованных сеточных интерфейсов. Среди предлагаемых в литературе методов согласования существует ряд алгоритмов, осуществляющих модификацию исходной сетки в области несогласованных интерфейсов [2, 3]. Однако с точки зрения вычислительной эффективности наиболее привлекательными являются консервативные методы, основанные на интерполяции величин в области несогласованных интерфейсов [4,5]. Такие методы не модифицируют и не перестраивают исходную сетку на интерфейсах, что значительно упрощает их программную реализацию.

В данной работе представлено описание метода, позволяющего учитывать особенности решения уравнений Навье-Стокса при моделировании течений вязкой несжимаемой жидкости в области смежных границ несогласованных произвольных неструктурированных фрагментов сетки. В основу метода положен принцип работы метода GGI (General Grid Interface) [5]. В отличие от метода GGI, в представленном методе консервативность интерфейса гарантируется за счет формирования идентичного набора виртуальных граней на смежных несогласованных интерфейсах. Предложен способ интерполяции величин на гранях несогласованных интерфейсов основанный на

методе усреднения по площади. Полученный метод позволяет объединять смежные несогласованные фрагменты сетки в единую расчетную область, учитывающую связь смежных ячеек через виртуальные грани, что позволяет сократить время решения задачи. Численные методы, представленные в данной работе, реализованы на базе отечественного пакета программ ЛОГОС [6]. Работоспособность представленного метода демонстрируется на трехмерном расчете задачи турбулентного течения в круглой трубе с внезапным сужением сеточная модель которой состоит из несогласованных фрагментов неструктурированной сетки. Результаты расчетов сравниваются с аналогичными расчетами на согласованной сетке.

Метод связывания несогласованных интерфейсов

Рассматривается математическая модель для решения задач течений вязкой несжимаемой жидкости, основанная на системе уравнений Навье-Стокса. Система уравнений решается путем численного интегрирования на конечно-объемной расчетной сетке. Дискретизацию дифференциальных уравнений методом конечных объемов рассмотрим на примере дискретизации уравнения переноса скалярной величины φ [7]:

$$\rho \frac{\partial \varphi}{\partial t} + \nabla \cdot (\rho \mathbf{u} \varphi) - \nabla \cdot (\Gamma \nabla \varphi) = 0, \quad (1)$$

В общем случае полный вид дискретного аналога дифференциального уравнения переноса скалярной величины (2) для контрольного объема P имеет вид:

$$\rho_P \frac{\varphi_P^n - \varphi_P^{n-1}}{\Delta t} V_P + C_P + D_P = 0, \quad (2)$$

где ρ_P - плотность в центре контрольного объема P , φ_P^n - скалярная величина в центре контрольного объема P на текущем временном шаге, φ_P^{n-1} - скалярная величина в центре контрольного объема P на предыдущем временном шаге, $C_P = \nabla \cdot (\rho_P \mathbf{u}_P \varphi_P)$ - конвективное слагаемое для контрольного объема P , $D_P = \nabla \cdot (\Gamma_P \nabla \varphi_P)$ - диффузионное слагаемое для контрольного объема P .

Приведенный алгоритм работы метода рассмотрим на примере дискретных аналогов конвективного и диффузионного слагаемых уравнения переноса при его дискретизации методом конечных объемов.

Рассмотрим расчетную область, состоящую из двух областей A и B (Рисунок 1), смежные границы которых не соответствуют друг другу по количеству и расположению узлов, образуя пару несогласованных интерфейсов. Для связывания несогласованных интерфейсов исходные грани заменяются на набор виртуальных граней. Геометрические параметры виртуальных граней определяются в результате последовательно проецирования исходных граней области A на все исходные грани области B . Полученный набор виртуальных граней образует виртуальный интерфейс.

Дискретный аналог конвективной составляющей для контрольного объема P , лежащего на несогласованном интерфейсе (Рисунок 1) принимает вид:

$$C_P = \sum_{f=face(P_{inner}, P_{bound})} \rho_f \varphi_f (\mathbf{u}_f \cdot \mathbf{S}_f) + \sum_{f=face(P_{interface})} \sum_v \rho_f^v \varphi_f^v (\mathbf{u}_f^v \cdot \mathbf{S}_f^v), \quad (3)$$

где $\rho_f^v, \varphi_f^v, \mathbf{u}_f^v$ - плотность, скалярная величина и скорость на виртуальной грани v , образовавшейся на исходной грани f, \mathbf{S}_f^v - площадь виртуальной грани v , ориентированная по нормали к исходной грани f .

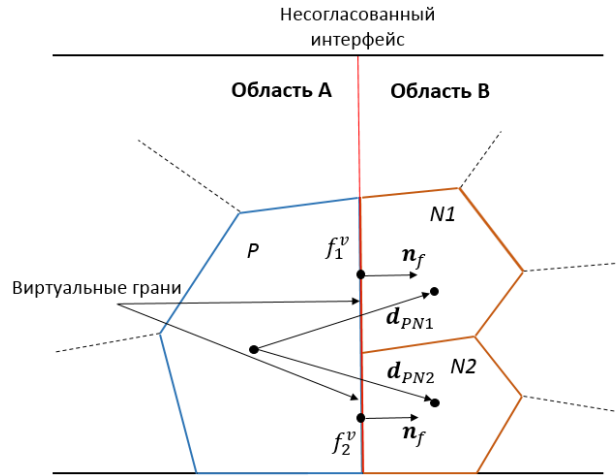


Рис. 1. Схематичный пример сетки с несогласованным сеточным интерфейсом

Дискретный аналог для диффузионной составляющей для контрольного объема Р, лежащего на несогласованном интерфейсе (Рисунок 1) примет вид:

$$D_P = \sum_{f=face(P_{inner}, P_{bound})} \Gamma_f (\nabla \varphi_f \cdot \mathbf{S}_f) + \sum_{f=face(P_{interface})} \sum_v \Gamma_f^v (\nabla \varphi_f^v \cdot \mathbf{S}_f^v), \quad (4)$$

где Γ_f^v , $\nabla \varphi_f^v$ - коэффициент диффузии и величина градиента скалярной величины на виртуальной грани V , образовавшейся на исходной грани f , \mathbf{S}_f^v - площадь виртуальной грани V , ориентированной по нормали к исходной грани f .

Все значения неизвестных величин ρ_f^v , φ_f^v , \mathbf{u}_f^v и Γ_f^v относятся к центру виртуальной грани V , образовавшейся на исходной грани f , и могут быть вычислены по любой из известных схем [9]. Значение градиента $\nabla \varphi_f^v$ в центре виртуальной грани V примем равным:

$$\nabla \varphi_f^v = \lambda_f^v \nabla \varphi_P + (1 - \lambda_f^v) \nabla \varphi_N, \quad (5)$$

где λ_f^v - интерполяционный фактор для виртуальной грани V , образовавшейся на интерфейсной грани f , $\nabla \varphi_P$, $\nabla \varphi_N$ - значения градиентов в центрах контрольных объемов Р и N соответственно.

В свою очередь значение градиента $\nabla \varphi_P$ в центре контрольного объема Р, вычисляется по алгоритму Грина-Гаусса [8]:

$$\nabla \varphi_P = \frac{1}{V_P} \left(\sum_{f=face(P_{inner}, P_{bound})} \varphi_f \mathbf{S}_f + \sum_{f=face(P_{interface})} \varphi_f^{av} \mathbf{S}_f \right), \quad (6)$$

где $\varphi_f^{av} = \frac{\sum_v \varphi_f^v \mathbf{S}_f}{\sum_v \mathbf{S}_f}$ - осредненное значение скалярной величины на интерфейсной грани f ,

входящей в контрольный объем Р.

В результате описанных выше преобразований, в дискретный аналог уравнения переноса (2) для каждого контрольного объема Р, лежащего на несогласованном интерфейсе (Рисунок 1) добавляются слагаемые, учитывающие связь смежных ячеек через виртуальные грани. Такой подход позволяет объединять смежные несогласованные

фрагменты сетки в единую расчетную область с учетом связей смежных ячеек через виртуальные грани.

Численный эксперимент

Рассмотрим работу метода на примере численного моделирования турбулентного течения несжимаемой жидкости в круглой трубе с внезапным сужением [9]. Геометрия расчетной модели и расположение контрольных сечений показанная на рисунке 2.

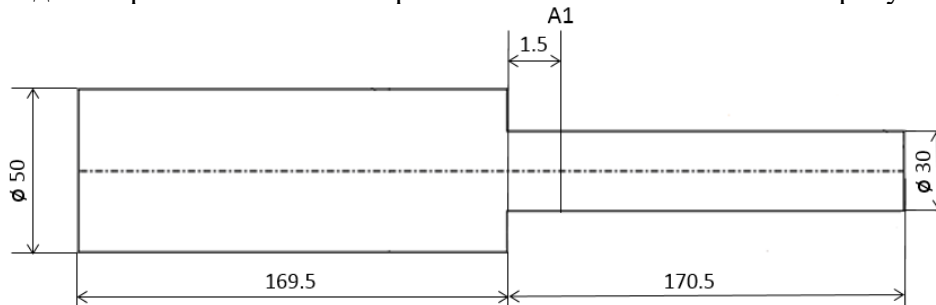


Рис. 2. Геометрия модели. Размеры указаны в мм

Для проверки работы представленного метода выполнены расчеты на согласованной и на несогласованной сеточных моделях. Согласованная сеточная модель получена путем сшивки фрагментов несогласованной модели, выполненной с помощью препостпроцессора пакета программ ЛОГОС.

Для расчетных моделей определяющими параметрами являются: плотность жидкости, равная $998,2 \text{ кг/м}^3$ и динамическая вязкость, равная $0,00101 \text{ кг/(м}\cdot\text{с)}$. Для моделирования турбулентного перемешивания используется k-Omega, SST Low модель турбулентности с параметрами: интенсивность турбулентности равным $0,05\%$ и длиной перемешивания равной $0,001 \text{ м}$. Для аппроксимации конвективных слагаемых в расчетах используется противопоточная схема первого порядка точности UD [9]. Внешние стенки заданы как непроницаемые жесткие стенки с прилипанием. Скорость потока на входе $V = 0,4227 \text{ м/с}$. На выходе задано давление $P = 0 \text{ Па}$. Для согласования решения в модели состоящей из несогласованных областей использованы интерфейсы, реализованные в пакете программ ЛОГОС соответствии с вышеописанным методом. Расчеты задач выполнены в параллельном режиме на 80 процессорах в стационарной постановке до сходимости по массе порядка 10^{-6} кг (Рисунок 3).

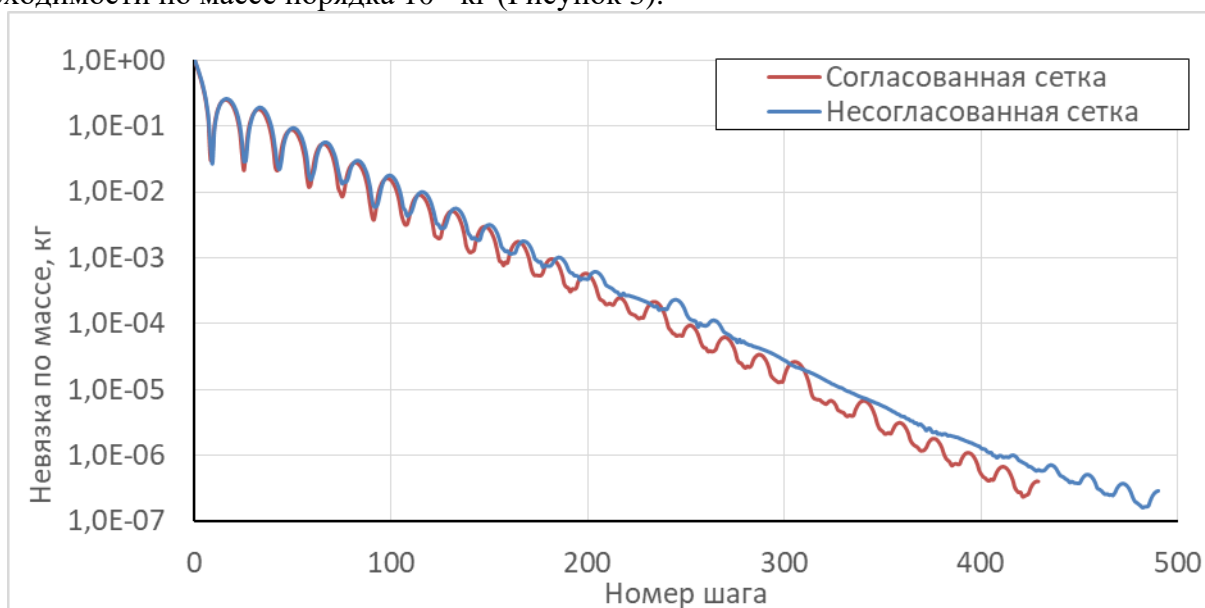


Рис. 3. График сходимости

При анализе результатов сравниваются профили скорости (Рисунок 4) и давления (Рисунок 5) вдоль горизонтальной линии, проходящей через середину канала, а также профили скорости (Рисунок 6) и давления (Рисунок 7) в поперечном сечении А1.

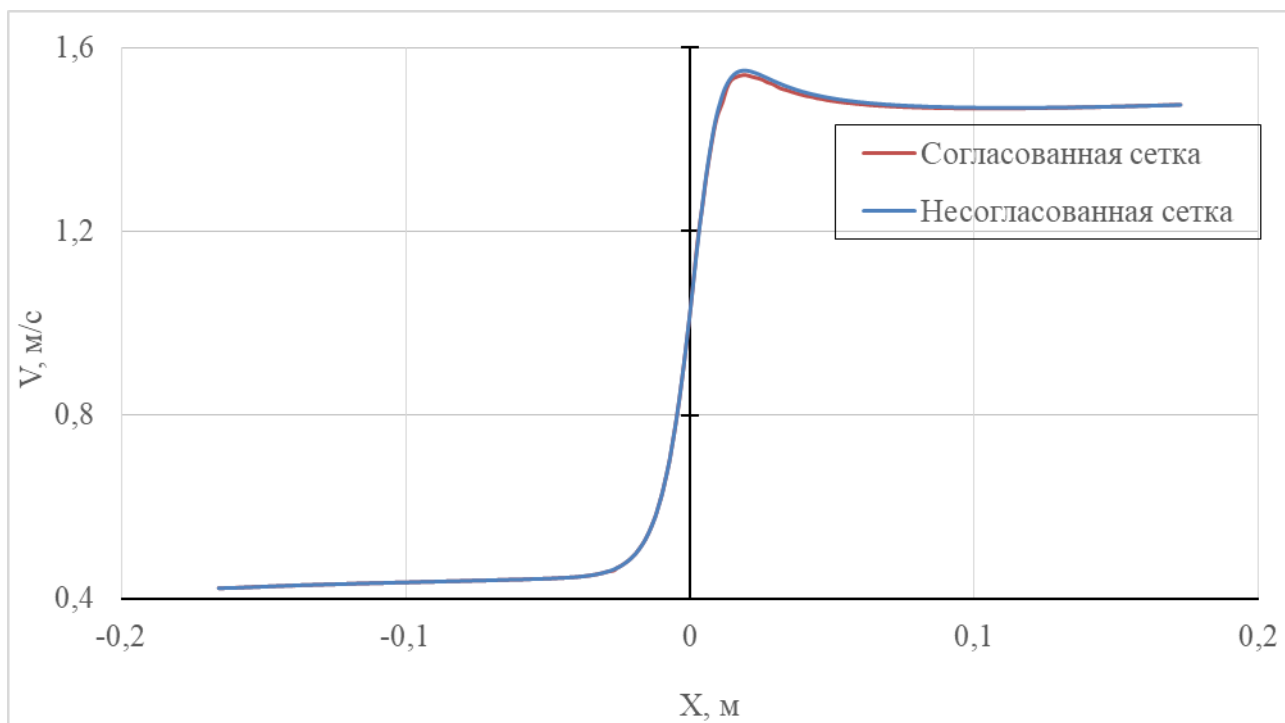


Рис. 4. Профиль скорости вдоль канала

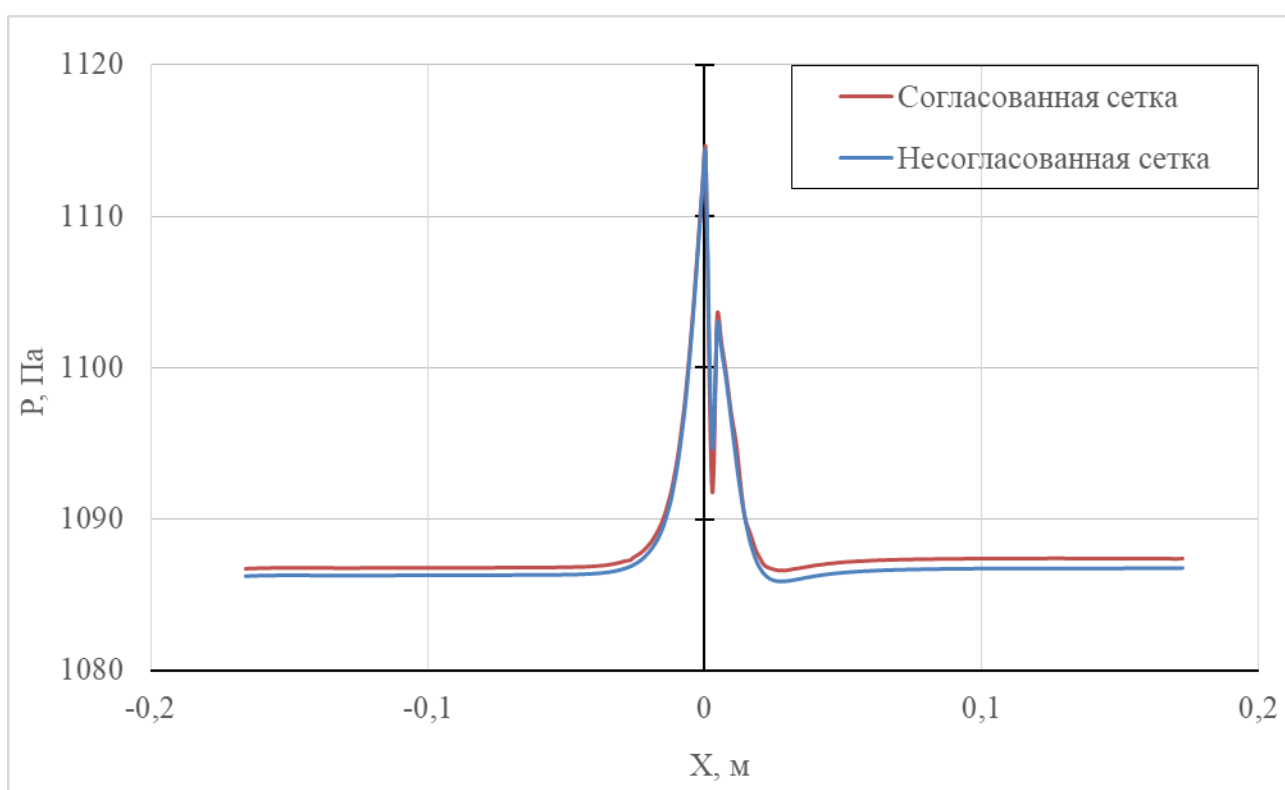


Рис. 5. Профиль давления вдоль канала

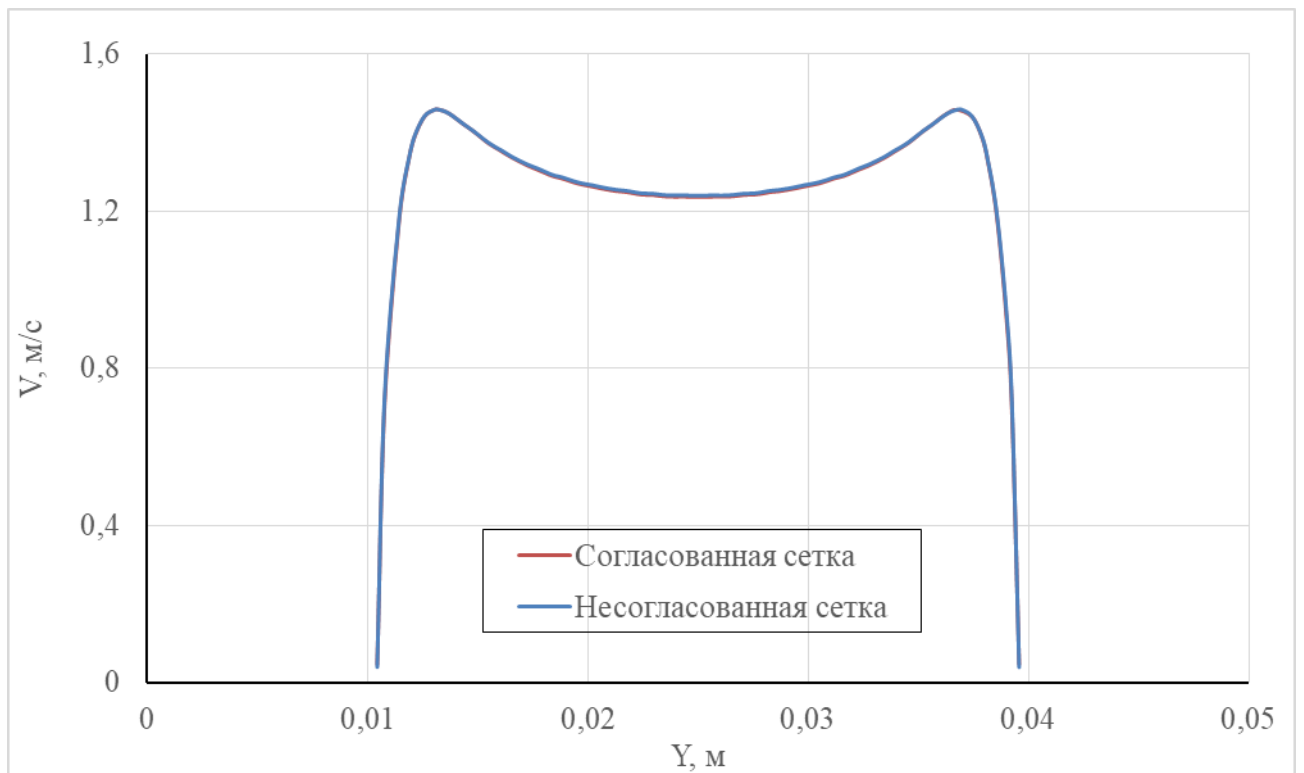


Рис. 6. Профиль скорости в сечении A1

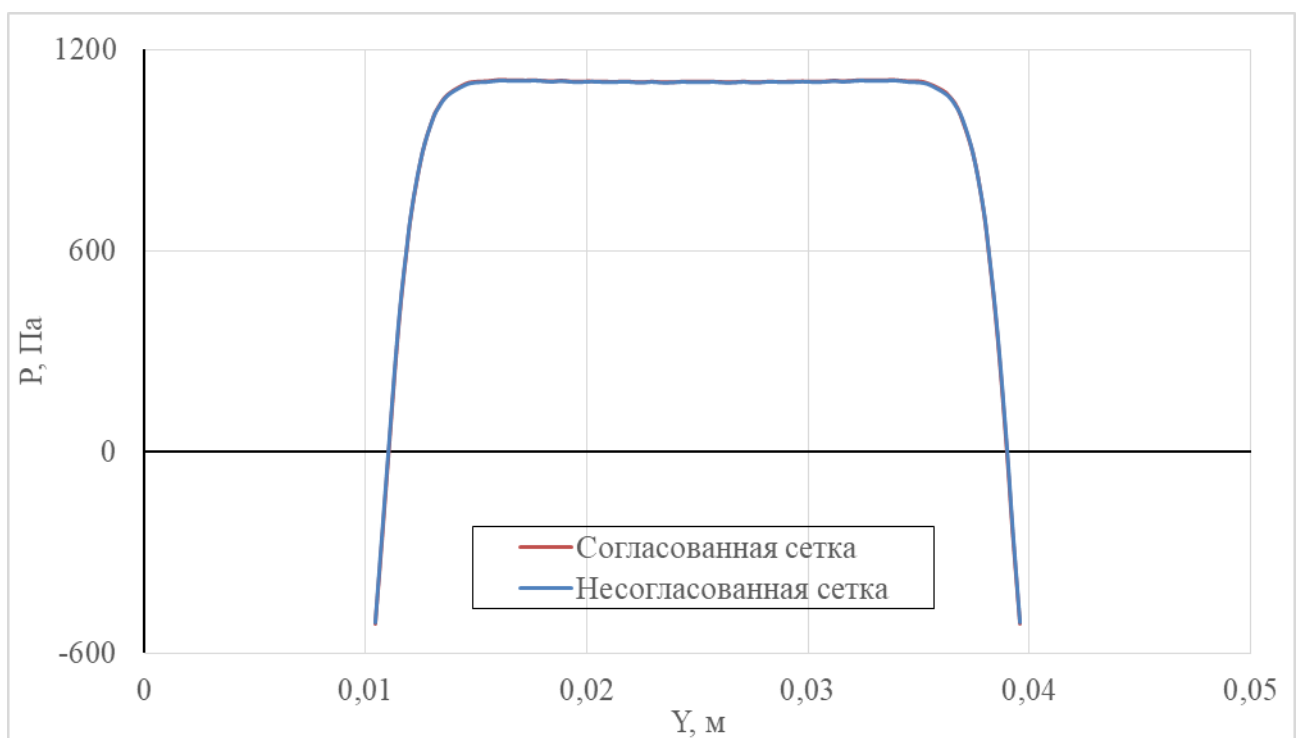


Рис. 7. Профиль давления в сечении A1

Полученные профили скорости и давления в продольном и поперечном сечениях показывают хорошее согласие результатов, полученных при расчете на согласованной и несогласованной сеточных моделях. Из графиков сходимости видно, что наличие несогласованного интерфейса несущественно влияет на скорость сходимости итерационного процесса. Полученные результаты расчетов демонстрируют эффективность разработанного метода.

Заключение

В статье представлено описание численного метода реализации несогласованного сеточного интерфейса. Метод позволяет объединять смежные несогласованные фрагменты неструктурированной сетки в единую расчетную область, с учетом связей смежных ячеек через виртуальные грани. Предложен способ интерполяции величин на границах несогласованных интерфейсов основанный на методе усреднения по площади.

Представлены результаты решения задачи турбулентного течения в круглой трубе с внезапным сужением с применением данного метода. Проведено сравнение профилей полей скорости и давления вдоль канала и в поперечном сечении на сеточных моделях, содержащих согласованные и несогласованные сетки. Показано что наличие несогласованных интерфейсов не оказывает существенного влияния на картину течения и скорость сходимости решения.

Библиографический список

1. Погосян М.А., Савельевских Е.П., Стрелец Д.Ю., Корнев А.В., Шагалиев Р.М., Козелков А.С., Использование отечественных суперкомпьютерных технологий при проектировании новых образцов авиационной техники // *Авиационная промышленность*. –2013.–№3– С. 3-7.
2. Qin N., Carnie G., Le Moigne A., Liu X., Shahpar S. Buffer layer method for linking two non-matching multi-block structured grids //47th AIAA Aerospace Sciences Meeting Including the New Horizons Forum and Aerospace Exposition: Orlando, January. – 2009.– P.5–8.
3. Wang Y., Qin N., Carnie G., Shahpar S., Zipper layer method for linking two dissimilar structured meshes // *J. Comput. Phys.*– 2013. – Vol. 255.– P.130–148.
4. Zhang Y., Chen H., Fu S., Improvement to patched grid technique with high-order conservative remapping method // *J. Aircr.* – 2011. – Vol. 48, no. 3.– P. 884–893.
6. Beaudoin Martin, Jasak Hrvoje Development of Generalized Grid Interface for Turbomachinery simulation with OpenFOAM / *Open Source CFD International Conference: Berlin,2008.*
1. Козелков А.С., Шагалиев Р.М., Курулин В.В., Ялозо А.В., Лашкин С.В. Исследование потенциала суперкомпьютеров для масштабируемого численного моделирования задач гидродинамики в промышленных приложениях // *Вычислительная математика и математическая физика*. – 2016. –Т. 56. – № 8. – С.1524–1535.
7. Лашкин С.В., Козелков А.С., Мелешкина Д.П., Ялозо А.В., Тарасова Н.В. Моделирование течений вязкой несжимаемой жидкости разделенным и совмещенным алгоритмом типа SIMPLE // *Математическое моделирование*. – 2016. – Т. 28. – № 6. – С. 64-76.
8. Ferziger J.H., Peric M. *Computational Method for Fluid Dynamics*. – Springer-Verlag, – 2002.
9. Идельчик И.Е. *Справочник по гидравлическим сопротивлениям*. - М.: Машиностроение, 1992.

A. V. Korotkov^{1,2}

**THREE-DIMENSIONAL NUMERICAL SIMULATIONS OF FLUID DYNAMICS
PROBLEMS ON GRIDS WITH UNMATCHED GRID INTERFACES**

*¹Federal State Unitary Enterprise “Russian Federal Nuclear Center All-Russian
Research Institute of Experimental Physics”*

²Nizhny Novgorod State Technical University n. a. R. E. Alekseev

This paper describes a numerical method, which considers specific computational fluid dynamics aspects of viscous incompressible flow simulations in the vicinity of interfaces between unmatched grid fragments. An example implementation of the method is presented for the case of the finite-volume approximation of the Navier-Stokes equations. The method is based on the GGI (General Grid Interface) principle, which does not require initial grid modification and involves conservative interpolation of fluxes at unmatched interfaces. The paper reports simulation results for a turbulent flow in a circular tube with an abrupt reduction in diameter.

Keywords: General Grid Interface, LOGOS, unmatched grid interface.

ОСОБЕННОСТИ МОДЕЛИРОВАНИЯ МНОГОКОМПОНЕНТНЫХ ТЕЧЕНИЙ НА ПРОИЗВОЛЬНЫХ НЕСТРУКТУРИРОВАННЫХ СЕТКАХ

¹Федеральное государственное унитарное предприятие «Российский Федеральный Ядерный Центр Всероссийский научно-исследовательский институт экспериментальной физики» (ФГУП «РФЯЦ - ВНИИЭФ»), г. Саров

²Федеральное государственное бюджетное образовательное учреждение высшего образования «Национальный исследовательский Томский государственный университет», г. Томск.

В статье рассматриваются особенности моделирования низкоскоростных многокомпонентных течений на произвольных неструктурированных сетках. Методика моделирования основана на алгоритме SIMPLE, предназначенном для численного моделирования несжимаемых и слабосжимаемых течений. Приведены результаты оценки диссипативности численных схем аппроксимации конвективных потоков для численного решения уравнения переноса концентраций на сетках, включая блочно-структурированные, а также произвольные неструктурированные тетраэдральные и полиэдральные. Оценка диссипативности осуществляется для схем первого порядка UD и гибридной схемы CD+0.1UD второго порядка точности.

Ключевые слова: многокомпонентные течения, неструктурированные сетки, уравнения переноса концентраций, уравнение Навье-Стокса.

Введение

Применение численных методов решения уравнений Навье-Стокса для расчета многокомпонентных течений сводится на практике в основном к расчетам на блочно-структурированных сетках с применением эффективных численных схем высокого порядка точности [1,2]. Однако при решении прикладных задач целесообразно использовать неструктурированные расчетные сетки, состоящие из многогранников произвольной формы с выделением призматических пограничных слоев для корректного разрешения турбулентного пограничного слоя [3], который влияет на течение в целом. Применение такого типа сеток позволяет моделировать задачу в областях со сложной геометрической конфигурацией в обозримые временные сроки. Однако, использование произвольной неструктурированной сетки приводит к ряду трудностей, одна из которых заключается в том, что на неструктурированной сетке заметно сужается круг доступных схем повышенного порядка точности для дискретизации конвективных потоков. От свойств выбранной схемы дискретизации напрямую зависит качество моделирования: схема должна иметь малую диссипацию и обеспечивать устойчивость счета [3]. Расчету на произвольных неструктурированных сетках посвящено много работ (см., например, [4,5]). Однако во всех перечисленных исследованиях не приводятся особенности при моделировании многокомпонентных течений, влияние используемых численных схем и их применимость при решении практических задач.

В настоящей работе проводится исследование выбранных схем аппроксимации конвективных потоков, предназначенных для расчета несжимаемых на неструктурированных сетках для моделирования низкоскоростных многокомпонентных течений. На примере решения задачи о впрыске струи пропана в воздушный поток оценивается влияние схем дискретизации конвективного потока в уравнении переноса концентраций.

Основные уравнения и метод расчёта многокомпонентных течений

Для моделирования низкоскоростных многокомпонентных слабосжимаемых течений применяют систему уравнения Навье-Стокса осредненную совместно с уравнениями переноса концентраций на основе закона диффузии Фика, которая имеет вид [6]:

$$\begin{cases} \frac{\partial \rho}{\partial \tau} + \nabla * (\rho \vec{u}) = 0, \\ \frac{\partial(\rho \vec{u})}{\partial \tau} + \rho(\vec{u} * \nabla) \vec{u} = -\nabla p + \mu \nabla^2 \vec{u} + \frac{1}{3} \mu \nabla(\nabla * \vec{u}), \\ C_p \frac{\partial(\rho T)}{\partial \tau} + \nabla * (\rho \vec{u} T) = -\nabla * (\lambda \nabla T) + \nabla(\vec{u} * \tau) + \frac{\partial \rho}{\partial \tau} + S_\tau, \\ \frac{\partial \rho c_i}{\partial \tau} + \nabla(\rho \vec{u} c_i + q_{c_i}) = S_{c_i}, i = 1 \dots n - 1 \end{cases} \quad (1)$$

Здесь \vec{u} - вектор скорости, ρ - плотность среды, p - давление, T - температура, μ - молекулярная вязкость, C_p - удельная теплоемкость, λ - эффективный коэффициент теплопроводности среды, $\tau = \mu(\nabla \vec{u} + (\nabla \vec{u})^T) - \frac{2}{3} \delta \nabla * \vec{u}$ - тензор вязких напряжений.

Для расчета концентраций компонент рассматриваются их объемные доли:

$$c_i, \quad 0 \leq c_i \leq 1, \quad \sum_{i=1}^n c_i = 1.$$

Последняя компонента вычисляется явно, по концентрациям других компонент:

$$c_n = 1 - \sum_{i=1}^{n-1} c_i.$$

Диффузионный поток компоненты c_i выражается при помощи коэффициента диффузии по закону Фика:

$$\vec{q}_{c_i} = -D_i \nabla c_i, i = 1 \dots n - 1, \quad (2)$$

где D_i - коэффициент диффузии компоненты. Диффузионный поток так же может быть выражен заданием числа Шмидта:

$$\vec{q}_{c_i} = -\frac{\mu}{Sc} \nabla c_i \quad (3)$$

В турбулентном случае, необходимо учитывать так же турбулентную составляющую диффузионного потока концентрации:

$$\vec{q}_{c_i} = -\left(\frac{\mu}{Sc} + \frac{\mu_t}{Sc_t}\right) \nabla c_i, \quad (4)$$

где μ_t - турбулентная вязкость, вычисленная с помощью какой-либо модели турбулентности. Уравнения дополнены возможными источниками энергии S_T и концентраций S_{c_i} .

Для дискретизации основных уравнений используется классический метод SIMPLE [7] на конечно-объемной трехмерной неструктурированной сетке [8]. Для моделирования турбулентности выбрана RANS модель SST [6], как одна из наиболее популярных и применимых для относительно низкоскоростных течений.

Рассматриваются интерполяционные схемы, шаблон которых ограничен парой смежных ячеек, поскольку более широкий шаблон на неструктурированной сетке в общем случае не существует, или требует построения сетки с выделенным направлением. Среди таких схем выбраны: простейшая схема первого порядка UD и гибридная схема CD + 0.1UD, представляющая собой линейную комбинацию схемы высокого и низкого порядков, совмещающая преимущества повышенного порядка и устойчивости счета на неструктурированной сетке [3,5,6]. Представленные схемы имеют следующий вид. Схема Upwind Differences - схема первого порядка, устойчива на неструктурированных сетках, но обладает большой численной диффузией [3,5,6]:

$$c_{i,F,ud} = \begin{cases} c_{i,P}, & f > 0 \\ c_{i,N}, & f < 0 \end{cases} \quad (5)$$

где P и N - индексы ячеек по разные стороны от грани с индексом F . Гибридные схемы представляют собой линейную комбинацию схемы высокого и низкого порядков, что приводит к увеличению монотонности решения. Гибридную схему, например, для схемы CD, можно записать следующим образом:

$$c_{i,F,gibr} = \gamma c_{i,F,CD} + (1 - \gamma) c_{i,F,ud} \quad (6)$$

где γ – коэффициент смешения. В настоящей работе рассматриваются схемы с постоянным коэффициентом смешения во всей расчетной области.

Численные эксперименты

Для моделирования многокомпонентных течений приведем ниже полученные результаты по пакету программ ЛОГОС – российский продукт инженерного анализа, предназначенный для решения сопряженных трехмерных задач конвективного теплопереноса, аэродинамики и гидродинамики на параллельных ЭВМ [3,5,7,8]. Для оценки сеточной сходимости рассматриваемого метода проведем эксперимент: моделирование многокомпонентных течений на блочно-структурированных сетках (O – сетках) и моделирование многокомпонентных течений на произвольных неструктурированных сетках.

Исследования многокомпонентных течений проводились на задаче впрыска струи пропана в воздушный туннельный поток полной трубы [9] (рисунок 1). В рассматриваемой задаче поток впрыскивается через форсунку в воздушное пространство трубы.

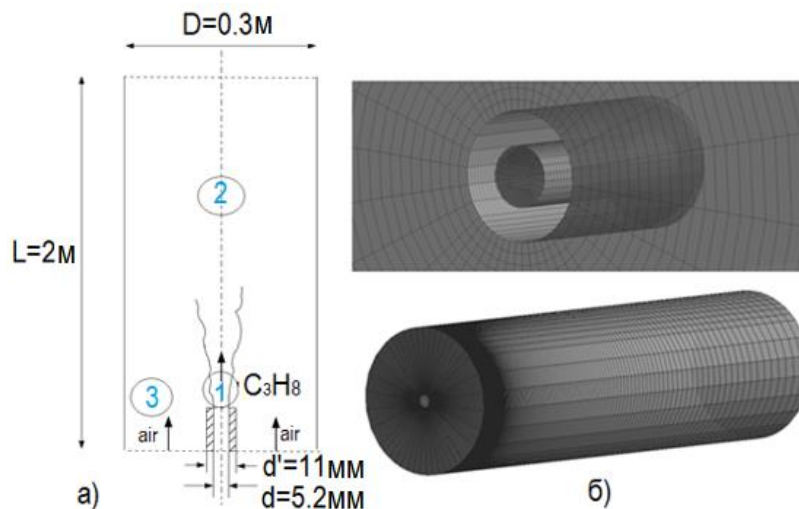


Рис. 1. Геометрия задачи (а) и вид расчетной блочно-структурированной O-сетки (б)

Исследуемое течение в трубе является турбулентным и содержит три компонента: пропан (C_3H_8 – 44.09 кг/моль; 7.95e-06 кг/мс; 1.91 кг/м³), кислород (O_2 – 31.9988 кг/моль; 1.919e-05 кг/мс; 1.2999 кг/м³) и азот (N_2 –28.0134 кг/моль; 1.663e-05 кг/мс; 1.138 кг/м³). Воздух, который входит в туннель свободен от C_3H_8 . Между компонентами нет никакой химической реакции, и поток адиабатный. Рассматривается развитие турбулентного течения пропана в потоке воздуха под действием силы тяжести. Воздух представлен смесью кислорода (O_2) и азота (N_2) в концентрации $C_{O_2} = 0,23$, $C_{N_2} = 0,77$. Струя пропана (C_3H_8) втекает со средней скоростью 53 м/с, число Рейнольдса $Re = 68000$. Длина туннеля составляет 2м, диаметр – 0.3м. Длина входного патрубка для пропана составляет 0,0165 м и была задана произвольно. Внешний диаметр патрубка равен 0.011м, внутренний диаметр – 0,0052м. На выходе из патрубка пропан начинает перемешиваться с воздухом. Для моделирования многокомпонентного газа используется уравнение состояния смеси

идеальных газов. В качестве коэффициента диффузии смеси использовалось число Шмидта, равное 0,7.

Для расчетов были построены 3 сетки для полной трубы - блочная O-сетка, тетраэдральная и полиэдральная сетки. Блочная O-сетка является базовой и состоит из 31 788 шестигранников со средним характерным размером 0,012 м. Произвольные неструктурированные тетраэдральная, и полиэдральная сетки строились с характерными размерами в 4 и 6 раз мельче - 0,004 м. и 0,002 м., согласно данным работ [3,5].

На рисунке 2 представлены распределения концентрации пропана, полученные с помощью схемы UD для всех трех сеток (слева направо). На рисунке 3 представлены распределения концентрации пропана, полученные с помощью схемы CD +0.1UD также для всех трех сеток (слева направо).

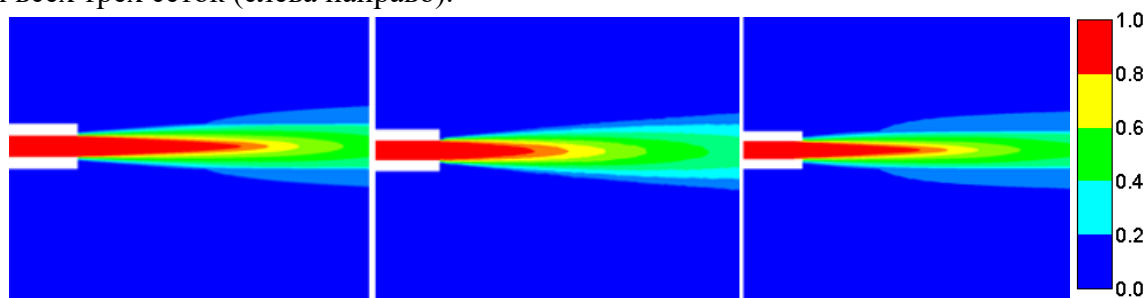


Рис. 2. Распределение концентрации пропана схемы UD

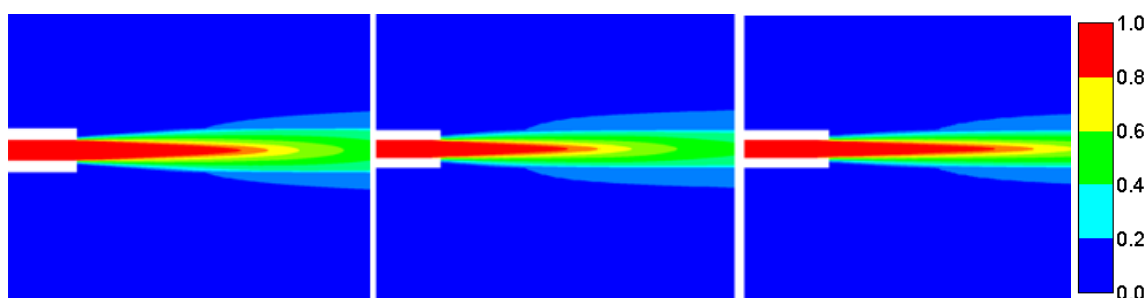


Рис. 3. Распределение концентрации пропана схемы CD +0.1UD

На рисунках 2 и 3 представлена анализируемая область смешения потоков, где градиенты концентраций максимальны. Видно, что ядро потока пропана, полученное на блочно-структурированной сетке наиболее выражено для схемы UD и практически совпадает по форме для гибридной схемы. Изолинии концентрации вне ядра потока имеют наибольшее отличие по форме для тетраэдральной сетки с учетом схемы UD, что объясняется повышенной диффузией из ядра потока. Для гибридной схемы видно отличие вне ядра для полиэдральной сетки, что объясняется большим количеством граней у каждой ячейки.

Для количественного сравнения **концентрации пропана** с экспериментальными данными на рисунке 4 приведено сравнение для подобласти начала смешения потоков согласно [9].

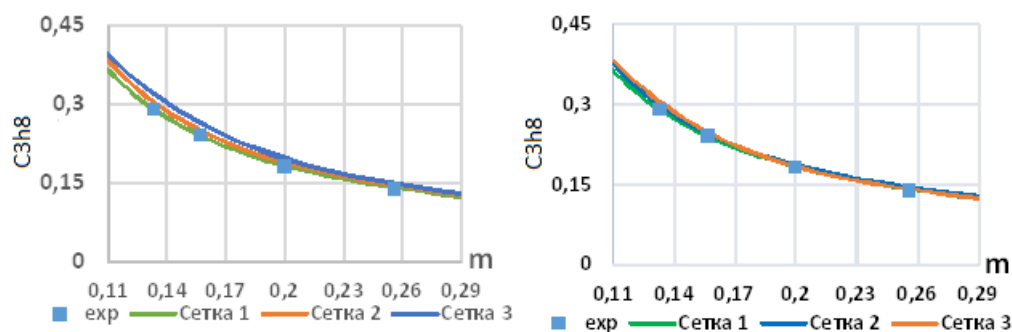


Рис. 4. Концентрация пропана на блочной (сетка 1), тетраэдральной (сетка 2) и полиэдральной (сетка 3) для схемы UD (слева) и CD +0.1UD (справа)

Для схемы первого порядка расчет на полиэдральных сетках демонстрирует предсказуемо большую численную диффузию. Расчет на тетраэдральной сетке несколько ближе к экспериментальным данным, наилучший результат получен на блочной сетке. Результаты на неструктурированных сетках сопоставимы с расчетом на блочной сетке только при соблюдении рекомендации о характерном размере ячеек [3,5].

Для гибридной схемы применение рекомендаций характерного размера неструктурированных ячеек позволяет достичь сопоставимой с блочной сеткой погрешностей, при устойчивой сходимости решения.

Заключение

В статье рассмотрено применение схемы повышенного порядка аппроксимации для расчета низкоскоростных многокомпонентных течений на произвольных неструктурированных сетках. По результатам расчетов приведены рекомендации по необходимому сеточному разрешению на основании полученных ячеек при построении и их характерному размеру, которые необходимы для получения приемлемых результатов.

Результаты расчетов, представленные в данной работе, показывают, что пакет программ ЛОГОС позволяет получать приемлемые результаты для решения задач многокомпонентности на неструктурированных сетках с характерным размером в 4 и 6 раз мельче [3,5] блочно-структурированных. Благодаря автоматической генерации неструктурированные сетки требуют значительно меньше трудозатрат при построении. В настоящее время данный подход не получил широкого применения на практике, однако имеет ряд преимуществ для задач со сложной геометрической конфигурацией. Как показано в данной работе, недостатки расчетов на неструктурированных сетках преодолимы для данного типа течений, но требуют дальнейшего исследования для течений, свойственных прикладным задачам распространения загрязнений.

Библиографический список

1. A Numerical Study of a Bluff-Body Stabilized Diffusion Flame. Part 2. Influence of Combustion Modeling and Finite-Rate Chemistry // a Department of Applied Mechanics, Thermodynamics and Fluid Dynamics, Norwegian University of Science and Technology, Trondheim, Norway Published online, 06 Apr 2007, P.191-217.
2. N. Peters Institut für Technische Mechanik RWTH Aachen // ERCOFTAC Summer School September 15-19, 1997 Aachen, Germany, P. 95.
3. Козелков А.С., Курулин В.В., Тятюшкина Е.С., Пучкова О.Л. Моделирование турбулентных течений вязкой несжимаемой жидкости на неструктурированных сетках с использованием модели отсоединенных вихрей // Математическое моделирование, 2014, Т. 26, № 8, С. 81-96.
4. Morton S.A., Forsythe J.R., Squires K.D., Wurtzler K.E. Assessment of unstructured grids for detached-eddy simulation of high Reynolds number separated flows // 8th ISGG Conf., Honolulu, June 2002.
5. Kozelkov A., Kurulin V., Emelyanov V., Tyatyushkina E., Volkov K. Comparison of convective flux discretization schemes in detached-eddy simulation of turbulent flows on unstructured meshes // Journal of Scientific Computing, 2016, V. 67, P. 176–191.
6. Ferziger J.H., Peric M., Computational Method for Fluid Dynamics // Springer-Verlag, New York, 2002, 423 p.
7. Козелков А.С., Дерюгин Ю.Н., Лашкин С.В., Силаев Д.П., Симонов П.Г., Тятюшкина Е.С. Реализация метода расчета вязкой несжимаемой жидкости с использованием многосеточного метода на основе алгоритма SIMPLE в пакете программ ЛОГОС // Вопросы атомной науки и техники, Математическое моделирование физических процессов, 2013, вып. 4, С. 44-56.
8. Волков К.Н., Емельянов В.Н., Козелков А.С. Методы ускорения газодинамических расчетов на неструктурированных сетках // Физматлит, 2014.
9. Fluent Validation Guide [Электронный ресурс]: Romeo.univ-reims // URL: <https://romeo.univ-reims.fr/documents/fluent/fluentValidationGuide.pdf>

K.B. Volodchenkova^{1,2}, R.R. Giniyatullin¹

**SPECIFIC ASPECTS OF MULTI-MATERIAL FLOW SIMULATIONS ON
ARBITRARY UNSTRUCTURED GRIDS**

*¹Federal State Unitary Enterprise “Russian Federal Nuclear Center All-Russian Research
Institute of Experimental Physics”*

²National Research Tomsk State University

The paper considers specific aspects of low-velocity multi-material flow simulations on arbitrary unstructured grids. The simulation technique is based on the SIMPLE method, which serves for numerical simulations of incompressible and low-compressible flows. The paper reports the results of dissipation analysis of convective flow approximation schemes as applied to numerical simulations of the volume fraction transfer equation on different types of grids, including block-structured and arbitrary unstructured tetrahedral and polyhedral grids. The dissipation analysis is performed for first-order UD schemes and the second-order CD+0.1UD hybrid scheme.

Keywords: multimaterial flows, unstructured grids, volume fraction transfer equations, Navier-Stokes equation.

НАВИГАЦИОННЫЙ КОМПЛЕКС БЕЗЭКИПАЖНОГО КОРАБЛЯ: ОПРЕДЕЛЕНИЕ РАССТОЯНИЯ ДО ОБЪЕКТА НА ВОДНОЙ ПОВЕРХНОСТИ С ПОМОЩЬЮ ЦИФРОВОЙ ВИДЕОКАМЕРЫ И ДАТЧИКОВ ПОЛОЖЕНИЯ

*Специальное конструкторское бюро средств автоматизации морских исследований ДВО
РАН, Южно-Сахалинск*

Предлагаемый метод обработки информации, получаемой с цифровой видеокамеры и датчиков положения в составе навигационного комплекса без экипажного корабля (БЭК), позволяет определять дистанцию, направление, а также относительную скорость и ускорение относительно объекта наблюдения. Предлагаемый метод может найти широкое применение в морской робототехнике для определения расстояния до объектов в надводной и подводной среде. Для обработки матрицы изображения используются математические операции, позволяющие получить сведения о движении относительно объекта наблюдения. Подобные системы можно комбинировать с машинным зрением, для увеличения автономности средств морской робототехники, а также для повышения безопасности управления наземными и наводными видами транспорта. Предлагаемая система обработки информации является модульной, позволяющей при объединении нескольких пространственных параметров расширить возможности системы управления БЭК.

Ключевые слова: машинное зрение, тригонометрические преобразования, аффинные преобразования, инерциальная навигация.

Введение

Системы определения расстояния до объектов наблюдения могут служить как вспомогательным инструментом для пилота (водителя) какого-либо транспорта, так и стать основной системой безопасности в навигации, предупреждая об опасности или автоматически корректируя движение. Целью работы является разработка специального программного обеспечения БЭК для определения расстояния до объекта наблюдения, внесенного в базу данных системы, при минимальном комплекте оборудования.

Известные Системы распознавания графических образов имеют следующие недостатки:

Сложность реализации – такие системы необходимо обучать перед эксплуатацией, используя набор признаков (Хаар, HOG, SURF и т.д.) [4]. Первоначальное обучение требует больших временных затрат как для подбора материалов и выделения у них интересующих признаков (если не имеются готовые библиотеки для задачи), так и время, потраченное на обучение, на обычном компьютере обучение может длиться часами, днями или даже неделями, в зависимости от потребностей;

Для решения подобных задач требуется достаточно мощная вычислительная система.

Предлагаемая система является модульной и может расширить свои функциональные возможности за счёт дополнительных датчиков, таких как гироскоп, акселерометр, модуль GPS, компас и т.д., например, акселерометр можно использовать для уточнения расстояния до объекта в надводной среде, во время качки, используя проекцию кажущегося ускорения для пересчёта значений.

Основным ядром системы является её аналитическая составляющая – в минимальной сборке достаточно откалибровать цифровую видео камеру (далее ЦВК) и ввести параметры ЦВК и объектов распознавания для того, чтобы система начала работать.

Алгоритм работы системы распознавания образов

Для распознавания образов целесообразно использовать наиболее отлаженную библиотеку с открытым исходным кодом OpenCV [5] (в .Net используется Emgu.CV), в которой для идентификации используется каскад Хаара. После обучения система способна довольно точно выполнять обнаружение всех объектов, которым обучил её пользователь.

Обнаруженные объекты интерпретируются как прямоугольник, у которого по умолчанию нам известна левая-верхняя точка, а также ширина и высота объекта относительно неё, из этих параметров необходимо найти центр основания объекта (рис. 1).

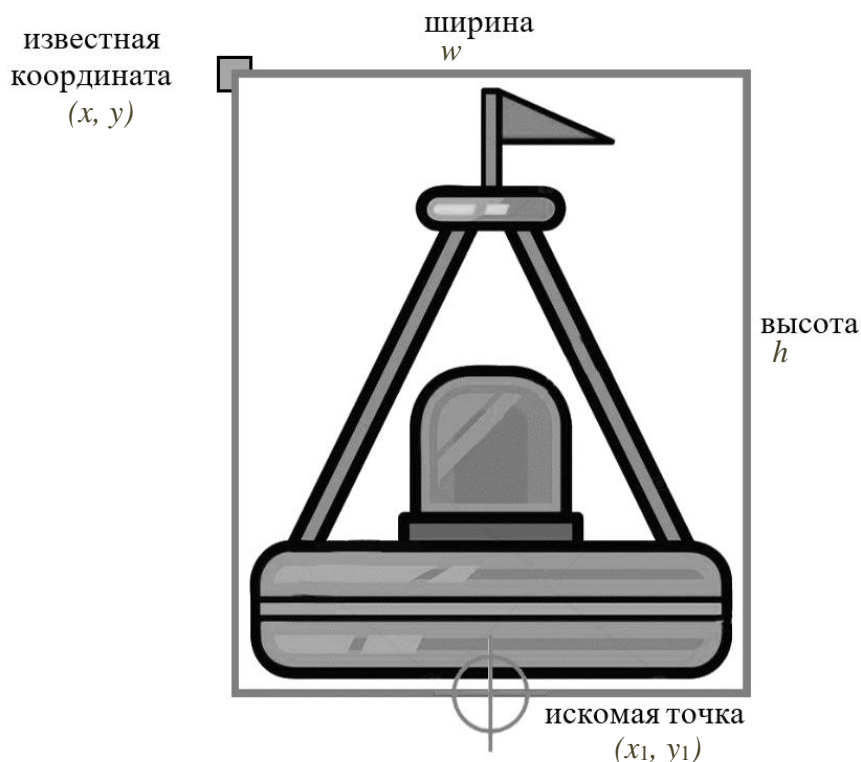


Рис. 1. Представление объекта в системе распознавания образов.

Изображения ЦВК можно представить в виде матрицы 1:

$$A = \begin{pmatrix} a_{11} & \dots & a_{x1} \\ \vdots & \ddots & \vdots \\ a_{1y} & \dots & a_{xy} \end{pmatrix}, \#(1)$$

размерностью $x \times y$, где для каждого обнаруженного объекта необходимо найти координату центра основания. Новые координаты обозначим как x_1 и y_1 , а высоту и ширину как h и w соответственно, где $x_1 = x + w/2$ и $y_1 = y + h$.

Зная внешние данные, либо внутренние мы можем узнать на каком расстоянии от нас находятся обнаруженные объекты [2, 3]. Для простоты понимая при работе с внешними данными, анализируется информация о видимых объектах и путём нахождения неких закономерностей, определяется расстояние. При работе с внутренними данными, нам известно наше положение в пространстве, в том числе расстояние от нас до горизонта, что в совокупности с данными ЦВК позволяет определить расстояние до объектов.

Метод тригонометрических функций с уточнением

Основная идея метода - нахождение расстояния при помощи тригонометрических расчетов и аффинных преобразований. В самой простой постановке задачи, решение имеет следующий вид $a = b * \tan \alpha$ (рис. 2). Здесь b – высота, α – угол до искомого объекта, a – искомое расстояние.

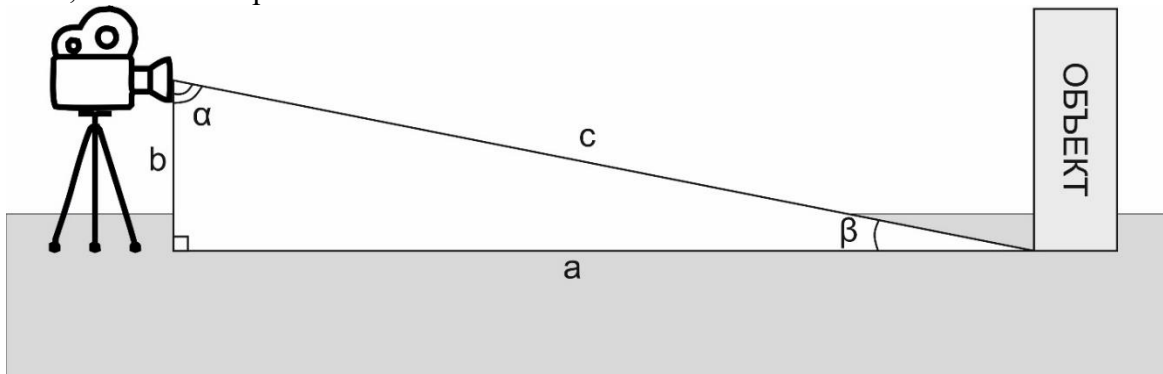


Рис. 2. Метод тригонометрических функций с уточнением

В частном случае, когда объекты находятся вдоль одной прямой:

$$d = h \times y / \tan \left(\frac{y}{\left(\frac{hm}{\alpha}\right)} + \left(90^\circ - \left(\beta + \frac{\alpha}{2}\right)\right) \right), \quad (2)$$

где d – искомое расстояние, м; h – высота на которой находится ЦВК, м; y – искомое значение по ординате; hm – высота матрицы; α – угол обзора ЦВК по вертикали, градус; β – угол под которым находится ЦВК относительно поверхности, градус (рис. 3).

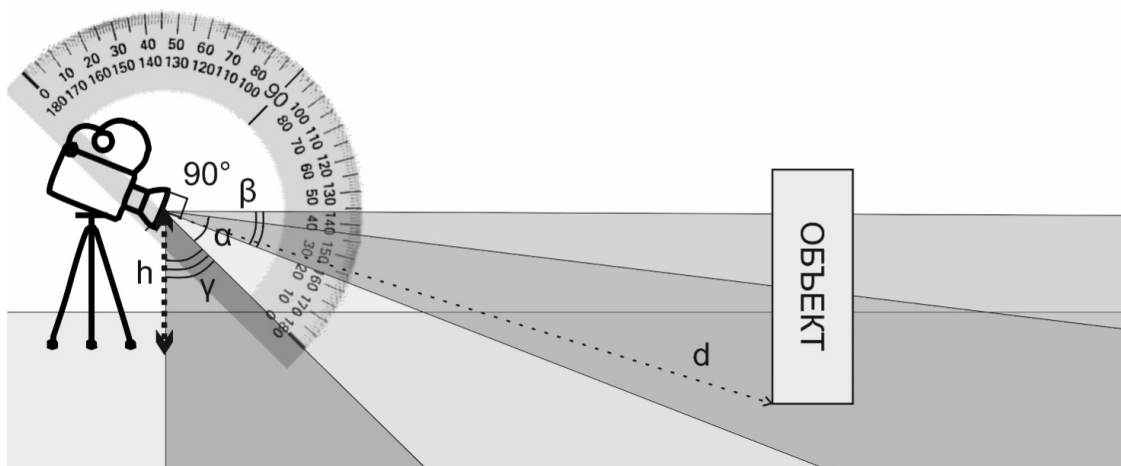


Рис. 3 Пояснения к формуле (3)

Рассмотрим вариант определения дистанции в трёхмерном пространстве. Расстояние до объекта, находящегося не на одной прямой с ЦВК (рис. 8):

$$ds = d / \cos \left(\frac{\frac{wm}{2} - x}{\frac{wm}{2 \times \alpha}} \right), \quad (3)$$

где ds – искомое расстояние, м; d – расстояние до объекта на одной прямой, м; wm – ширина матрицы; x – искомое значение по абсциссе; α – угол обзора ЦВК по горизонтали (абсциссе), градус.

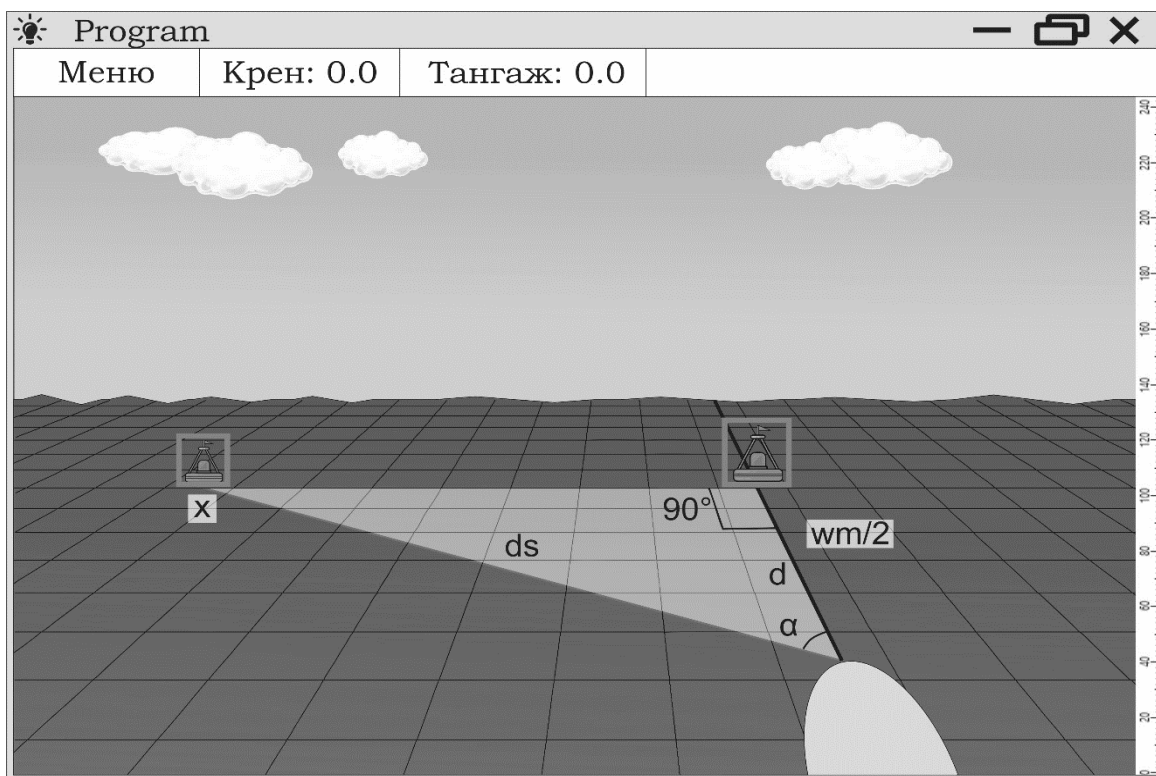


Рис. 4. Пояснения к определению расстояния по формуле (3)

Далее рассмотрим вариант определения расстояния, когда в условиях морского волнения изменяется тангаж.

При условии, когда ЦВК меняет свой тангаж, и объект всё так же попадает в объектив и идентифицируется, необходимо учесть эти отклонения в расчётах. Принимаем, что «точка объекта» — это середина его основания. Для того, чтобы отображаемое расстояние не изменилось, мы создаём новую виртуальную точку, которая будет находиться в том же месте, где она находилась бы при нулевом тангаже. Выразим это в виде формулы:

$$y = y_0 + \frac{hm}{\alpha} \times \beta, \# \quad (4)$$

где y – искомое значение по ординате; y_0 – начальное значение по ординате, при нулевом тангаже; hm – высота матрицы; α – угол обзора камеры по вертикали, градус; β – тангаж.

Для пересчёта координат объекта при изменении крена, мы используем аффинные преобразования, представим матрицу преобразований в виде отдельных формул корректировки для абсциссы и ординаты.

Рассмотрим поворот фигуры вокруг произвольной точки на произвольный угол с помощью матрицы. Чтобы произвести любое сложное преобразование, необходимо разложить его на базовые операции. Поворот фигуры или в данном случае точки вокруг другой произвольной точки (x_0, y_0) на произвольный угол α , состоит из трёх базовых операций:

- 1) Перенос точки на вектор $A(-x_0, -y_0)$ для совмещения точки (x_0, y_0) с началом координат.
- 2) Поворот точки на угол α .
- 3) Перенос точки на вектор $A'(x_0, y_0)$ для возвращения её в исходное положение.

Пусть мы хотим повернуть точку с координатами $A(x,y)$ вокруг точки $D(x_0,y_0)$ на угол α . Пусть P_s – матрица переноса точки на вектор $A(-x_0,-y_0)$, V_α – матрица поворота на угол α , P_s – матрица переноса точки на вектор $A'(x_0, y_0)$ [1].

$$\begin{aligned} (x \ y \ 1) \times P_{-s} \times V_\alpha \times P_s &= (x \ y \ 1) \times \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ -x_0 & -y_0 & 1 \end{pmatrix} \times \\ &\times \begin{pmatrix} \cos \alpha & \sin \alpha & 0 \\ -\sin \alpha & \cos \alpha & 0 \\ 0 & 0 & 1 \end{pmatrix} \times \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ x_0 & y_0 & 1 \end{pmatrix} = (x \ y \ 1) \times \\ &\times \begin{pmatrix} \cos \alpha & \sin \alpha & 0 \\ -\sin \alpha & \cos \alpha & 0 \\ \cos \alpha * x_0 - \sin \alpha * y_0 - x_0 & \sin \alpha * x_0 - \cos \alpha * y_0 - y_0 & 1 \end{pmatrix} \end{aligned} \quad (5)$$

Для удобства представления и адаптации матрицы под численные методы, представим её в виде формулы для x и y . Исходя из полученной матрицы:

$$x = x_0 + (x - x_0) * \cos(\alpha) - (y - y_0) * \sin(\alpha), \#(6)$$

$$y = y_0 + (y - y_0) * \cos(\alpha) + (x - x_0) * \sin(\alpha) \#(7)$$

Координаты (x,y) являются центром основания обнаруженного объекта, угол α является креном и определяется акселерометром. Для решения формулы нам не хватает координаты точки вращения. Обратимся к рисунку 4, где прямая расположенная перпендикулярно плоскости, на которой находится ЦВК и является точкой (x_0,y_0) . В данном примере точка вращения находится за пределами матрицы ЦВК (не попадет в объектив камеры), это вполне приемлемо, в таком случае координата будет иметь отрицательное значение. В общем виде можно представить:

$$x_0 = \frac{wm}{2}, \#(8)$$

$$y_0 = (90 - (\gamma + \alpha)) * c, \#(9)$$

где wm – ширина матрицы; γ – угол между перпендикуляром к поверхности и начальным углом обзора ЦВК, градус; α – угол обзора ЦВК, градус; c – отношение одного градуса угла к высоте пикселей.

Заключение

Для определения дистанции до объекта наилучшим методом является использование тригонометрических функций и аффинных преобразований. Этот метод не требует большого количества исходных данных и легко адаптируется под новое оборудование и условия эксплуатации.

Таким образом, система, имеющая лишь цифровую камеру и акселерометр, уже может определять расстояние до объектов. Если к системе добавить компас и датчик GPS, появится возможность определить геолокацию объекта и отобразить его на карте. Предлагаемая система может сильно удешевить и упростить разработку безэкипажных кораблей, систем безопасности (видеонаблюдения).

Представленные результаты получены при финансовой поддержке грантов Президента РФ МД-148.2020.5 и НШ-2485.2020.5.

Библиографический список

1. Демин А. Ю., Кудинов А. В. Компьютерная графика. Учеб. пособие // Томск: Том. политехн. ун-т., 2005. 160с.
2. Локтев Д.А., Алфимцев А.Н. Измерение расстояния до движущегося объекта с помощью комплексной системы видеомониторинга // Наука и инновации. 2013. № 11 (23). С. 4. doi: 10.18698/2308-6033-2013-11-996.
3. Локтев А.А., Локтев Д.А. Метод определения расстояния до объекта путем анализа размытия его изображения // Вестник МГСУ. 2015. № 6. С. 140-151.
4. Половинкин П.Н. Детекторы и дескрипторы ключевых точек. Алгоритмы классификации изображений. Задача детектирования объектов на изображениях и методы её решения // Учебный курс «Разработка мультимедийных приложений с использованием библиотек OpenCV и IPP», 2013. Нижний Новгород: Нижегородский государственный университет им. Н.И. Лобачевского. 38с.
5. Официальный сайт библиотеки OpenCV. <https://opencv.org/>.

Aleshin S.A., Belanov I.A., Leonenkov R.V., Sidorenko A.I.

NAVIGATION COMPLEX OF STEALTHLESS SHIP: DETERMINATION OF DISTANCE TO OBJECT ON WATER SURFACE USING DIGITAL VIDEO CAMERA AND POSITION SENSORS

Special Research Bureau for Automation of Marine Researches Far East Branch Russian Academy of Sciences, Yuzhno-Sakhalinsk

The proposed method for processing information received from a digital video camera and position sensors as part of a navigation complex without a crewed ship (BEC) allows determining the distance, direction, as well as the relative speed and acceleration relative to the object of observation. The proposed method can find wide application in marine robotics for determining the distance to objects in the surface and underwater environment. To process the image matrix, mathematical operations are used to obtain information about the movement relative to the observation object. Such systems can be combined with machine vision to increase the autonomy of marine robotics, as well as to improve the safety of control of land and water transport. The proposed information processing system is modular, allowing, by combining several spatial parameters, to expand the capabilities of the BEC control system.

Keywords: machine vision, trigonometric transformations, affine transformations, inertial navigation

ТЕСТОВЫЕ РАСЧЕТЫ РАСПРОСТРАНЕНИЯ ПОВЕРХНОСТНЫХ ВОЛН В РАМКАХ НЕЛИНЕЙНО-ДИСПЕРСИОННОЙ ТЕОРИИ ДЛИННЫХ ВОЛН

¹Нижегородский государственный технический университет им. Р.Е. Алексеева,

²Национальный исследовательский нижегородский государственный университет им. Н.И. Лобачевского,

³Институт прикладной физики РАН, Нижний Новгород

Проведены тестовые расчеты распространения волн на поверхности жидкости конечной глубины в рамках нелинейно-дисперсионной теории длинных волн. В расчетах использован вычислительный комплекс CLAWPACK-BOUSSCLAW, решающий двумерные уравнения Буссинеска. Исследуется в деталях влияние дисперсии на характеристики распространяющейся волны.

Ключевые слова: длинные волны, численный эксперимент, уравнения мелкой воды, дисперсия, уравнения типа Буссинеска

Введение

Теория мелкой воды активно применяется для описания катастрофических волн (цунами, штормовые нагоны) в прибрежной зоне. Достаточно указать несколько книг [1-4], где обсуждаются различные аспекты теории длинных волн. В наших расчетах динамики волн на мелкой воде мы неоднократно использовали эту теорию, и численно решали соответствующие уравнения с помощью вычислительного комплекса CLAWPACK [5]. Между тем, реальные морские волны не всегда являются достаточно длинными (по сравнению с глубиной), так что для них необходимо учитывать дисперсионные эффекты. Нелинейно-дисперсионная теория длинных волн также активно развивалась последние годы [6-8]. Выведенные уравнения с использованием малого параметра (отношение глубины к длине волны) получили название уравнений Буссинеска, и их главное достоинство заключается в двумерности уравнений, несмотря на трехмерность волнового движения. Существует большое разнообразие уравнений Буссинеска, связанных с аппроксимацией дисперсионного соотношения в длинноволновой области, а также соответствующих численных кодов. В частности, в [9] разработан вычислительный пакет CLAWPACK-BOUSSCLAW, позволяющий решать уравнения Буссинеска. В настоящей работе приведены результаты тестовых расчетов динамики нелинейных волн на воде с учетом дисперсии в сопоставлении с результатами, полученных в рамках теории мелкой воды, выполненных с помощью вычислительного комплекса CLAWPACK-BOUSSCLAW.

Математическая модель

Использовались две математические модели: классическая система нелинейных уравнений мелкой воды и уравнения типа Буссинеска.

Первая математическая модель представлена одномерной нелинейной системой уравнений мелкой воды (1), записанной в дивергентной форме и представляющей законы сохранения массы и импульса:

$$\begin{cases} \frac{\partial H}{\partial t} + \frac{\partial Hu}{\partial x} = 0, \\ \frac{\partial Hu}{\partial t} + \frac{\partial}{\partial x} \left(Hu^2 + \frac{1}{2} gH^2 \right) = gh \frac{dh}{dx}. \end{cases} \quad (1)$$

Здесь $H(x, t) = h(x) + \eta(x, t)$ – это полная глубина воды, $\eta(x, t)$ – смещение водной поверхности, $h(x)$ – невозмущенная глубина воды, $u(x, t)$ – горизонтальная скорость потока, g – гравитационная постоянная.

Используемые в расчетах уравнения типа Буссинеска имеют вид:

$$H_t + (Hu)_x = 0, \quad (2)$$

$$(1 - D)[(Hu)_t] + \left(Hu^2 + \frac{g}{2} H^2 \right)_x - gHh_x - Bgh^2(h\eta_x)_{xx} = 0. \quad (3)$$

Оператор D определяется с помощью фиктивной переменной $w(x, t)$ следующим образом:

$$D(w) = \left(B + \frac{1}{2} \right) h^2 w_{xx} - \frac{1}{6} h^3 \left(\frac{w}{h} \right)_{xx} \quad (4)$$

для любых $w(x, t)$. В уравнениях (3) и (4) B – дисперсионный параметр.

Мадсен и Соренсен в своей работе [10] использовали $B = 1/15$, при котором дисперсионное соотношение, вытекающее из уравнений Буссинеска, лучше соответствует точному соотношению. Когда $B = 0$, эта система уравнений типа Буссинеска приблизительно сводится к системе Перегринна [11], поскольку линейные дисперсионные соотношения идентичны. Однако, в отличие от уравнения количества движения в системе Перегринна, гидростатические части (4) записаны в консервативной форме. Кроме того, в дисперсионный член вносится некоторая нелинейность. Система уравнений (2), (3) и (4) является приближенной по параметрам нелинейности и дисперсии.

Все численные эксперименты проводились с помощью программного обеспечения CLAWPACK и его надстройки BOUSSCLAW (www.clawpack.org). Пакет CLAWPACK решает систему уравнений мелкой воды (1) методом конечных объемов [12]. Вычислительный пакет BOUSSCLAW использует гибридный метод для численного решения системы уравнений (2) и (3): методы конечных объемов и конечных разностей. В частности, используется метод конечного объема для нелинейной части уравнений и дискретизация конечных разностей с дробными шагами для дополнительных членов, таких как члены дисперсии, как стандартного, так и более высокого порядка.

Численные эксперименты

В рамках исследований динамики длинных волн, основанных на дисперсионной и бездисперсионной моделях, была проведена серия экспериментов по эволюции начального возмущения одиночного типа. Длина расчетной области выбрана 400 м, и на ее границах применялись условия свободного ухода волны из области (условия Зоммерфельда), они описаны в [13]. Глубина воды принималась равной 1 м. Шаг по пространству 0.15 м, шаг по времени настраивался автоматически для выполнения условия Куранта-Фридрихса-Леви [14].

Начальное возмущение задавалось в виде колоколообразного импульса солитоноподобной формы (рис. 1):

$$\eta_{in}(x, 0) = A_{ini} \cosh^{-2}[(x - x_0)/L], \quad (5)$$

причем его амплитуда могла быть как положительной, так и отрицательной. В наших расчетах характерная полудлина $L = 6$ м подобрана таким образом, чтобы длина начального импульса у подошвы составляла около 50 м. Пик падающего импульса располагался в точке $x_0 = 100$ м. Начальная скорость течения жидкости принималась равной нулю, поэтому начальный импульс расщеплялся две волны, которые затем

распространялись независимо. Амплитуда волны A_{init} варьируется от 0,001 до 0.5 м как для впадин, так и для гребней. В рамках линейной теории мелкой воды амплитуды свободных волн (A_0) была вдвое меньше начальной.

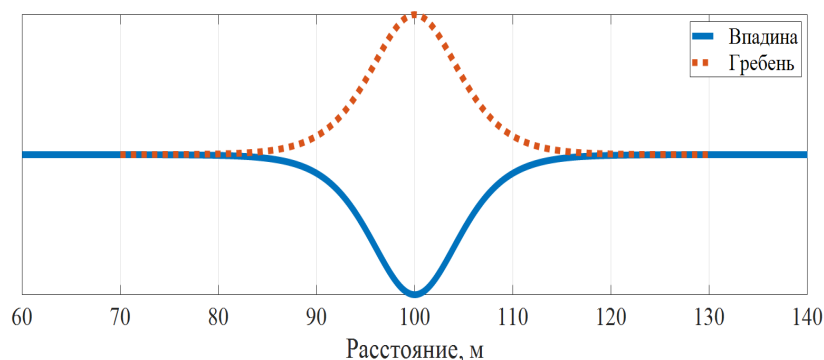


Рис. 1. Форма начального возмущения

Эффект дисперсии сначала исследовался для волн малой амплитуды ($A_0 = 1$ мм), для которых можно ожидать отсутствие нелинейных эффектов (рис. 2).

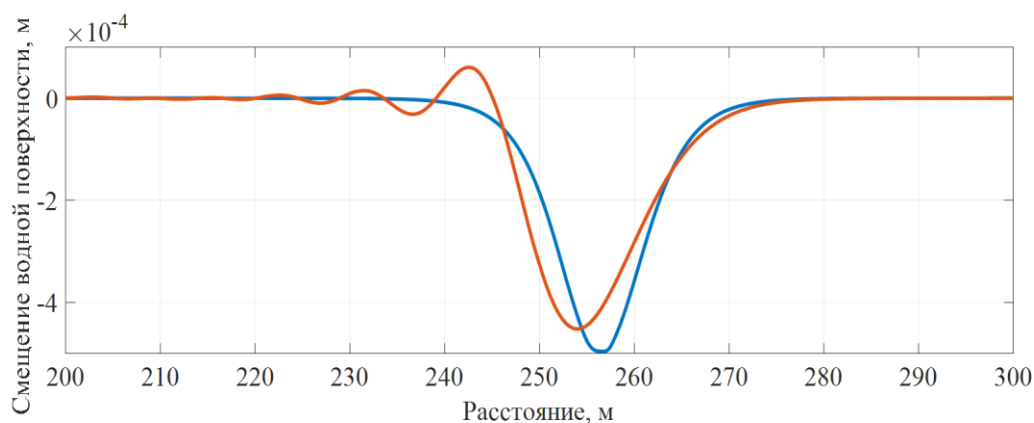


Рис. 2. Форма волны в момент времени $t=100$ с. Начальная амплитуда $A_0=0.001$ м. Синяя кривая соответствует бездисперсионной модели, красная – дисперсионной

Как и следует из теории, за головной волной появляются осцилляции при учете дисперсии, и волна развивается в известную волну Эйри. В момент времени 100 с максимальная амплитуда волны, рассчитанная с помощью дисперсионной модели, составляет 0.45 мм (за счет ухода энергии в дисперсионный хвост), в то время как в бездисперсионной модели она составляет 0.5 мм в соответствии с предсказаниями линейной теории.

При увеличении амплитуды волны до 0.1 м (Рис. 3) начинают сказываться нелинейные эффекты, приводящие в теории мелкой воды к образованию ударного фронта на заднем склоне и более сильному процессу образования дисперсионного хвоста в нелинейно-дисперсионной теории. Эти эффекты хорошо известны в теории, и наши результаты можно рассматривать как тестирование вычислительного пакета CLAWPACK.

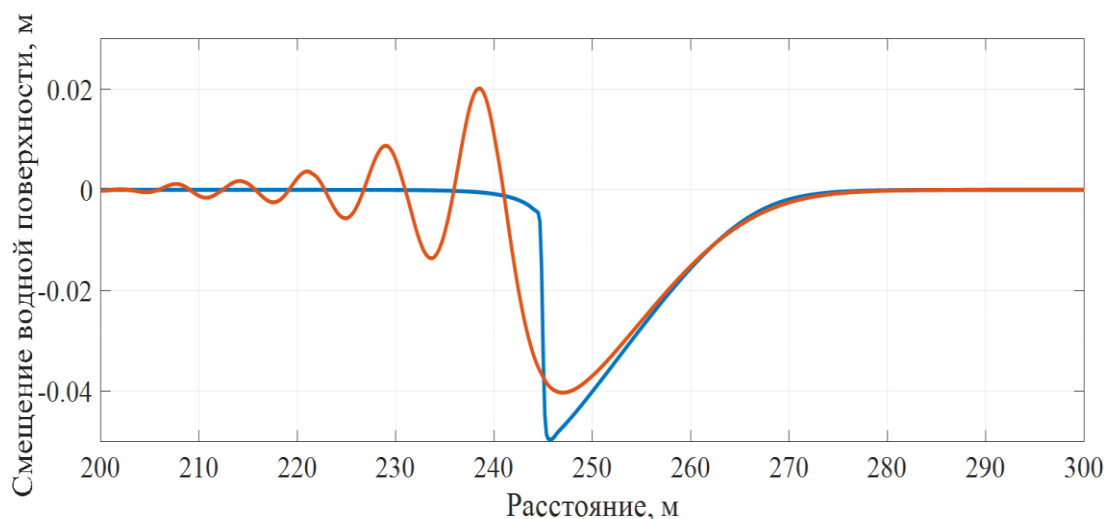


Рис. 3. Форма волны в момент времени $t=100$ с. Начальная амплитуда $A_0=0.1$ м. Синяя кривая соответствует бездисперсионной модели, красная – дисперсионной

При дальнейшем увеличении амплитуды впадины до 0.5 м (Рис. 4) нелинейные и дисперсионные эффекты становятся еще более заметными, ударная волна в рамках теории мелкой воды быстро затухает, а количество осцилляций в дисперсионном хвосте в рамках уравнений Буссинеска) сильно возрастает. Между тем вдали от основной волны наблюдаются нерегулярные колебания, видимые на рис. 4, их природа скорее всего связана с недостаточно точной реализацией граничных условий свободного ухода.

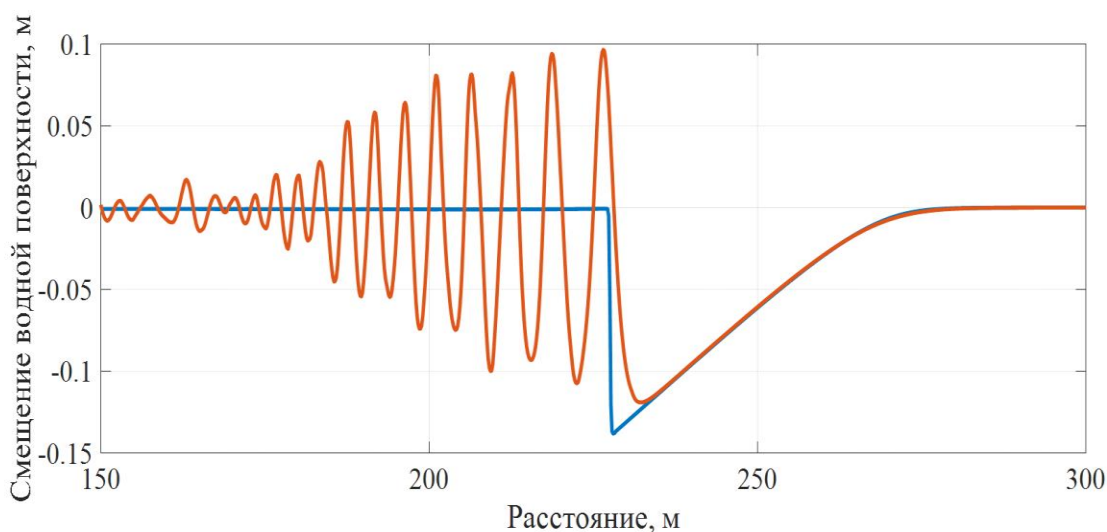


Рис. 4. Форма волны в момент времени $t=100$ с. Начальная амплитуда $A_0=0.5$ м. Синяя кривая соответствует бездисперсионной модели, красная – дисперсионной

Рассмотрим теперь случай распространения волнового гребня. При очень малых амплитудах ($A_0 = 0.1$ мм), как видно из рис. 5, мы имеем фактически зеркальное отражение рис. 2, как и ожидалось.

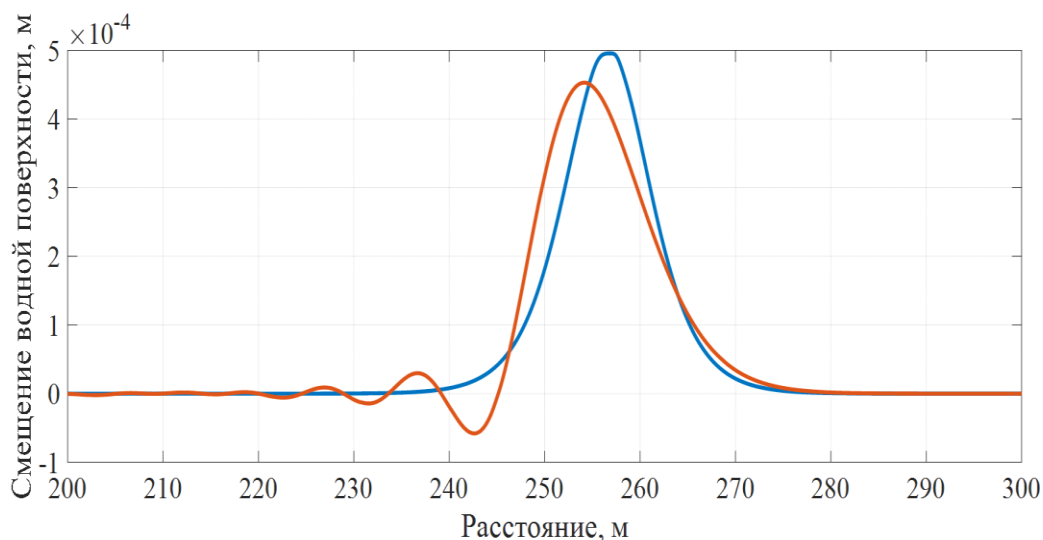


Рис. 5. Форма волны положительной полярности в момент времени $t=100$ с. Начальная амплитуда $A_0=0.001$ м. Синяя кривая соответствует бездисперсионной модели, красная – дисперсионной

С увеличением амплитуды волны начинается сильное усиление нелинейных и дисперсионных эффектов, приводящее к генерации уединенных волн (солитонов). Так, при начальной амплитуде 0.1 м (Рис. 6) с учетом дисперсии генерируется уединенная волна (солитон) практически без хвостов, в то время как без дисперсии формируется ударная волна.

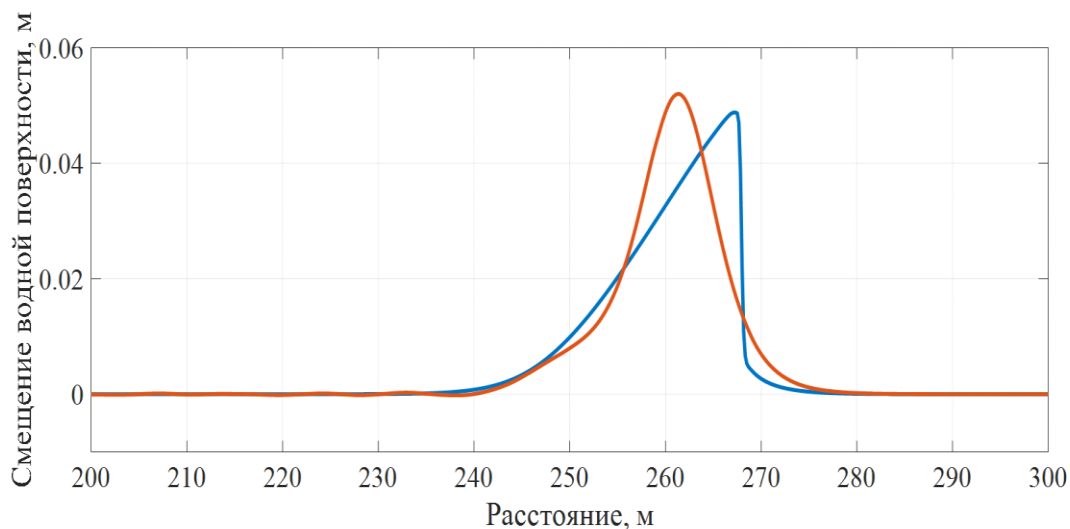


Рис. 6. Распространение гребня в момент времени $t=100$ с. Начальная амплитуда $A_0=0.1$ м. Синяя кривая соответствует бездисперсионной модели, красная – дисперсионной

При амплитуде 0.5 м появляются уже два солитона (рис. 7), хотя второй солитон еще не полностью сформировался.

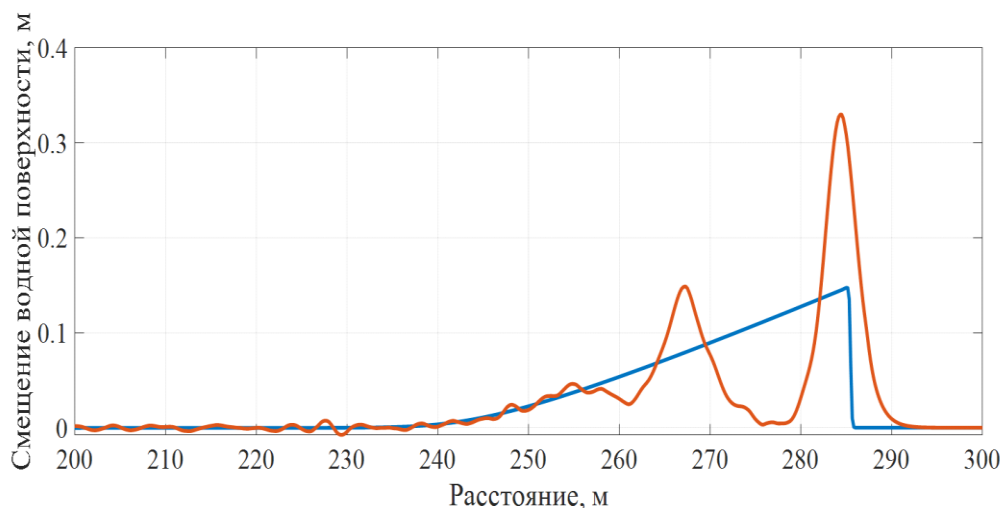


Рис. 7. Распространение гребня в момент времени $t=100$ с. Начальная амплитуда $A_0=0.5$ м. Синяя кривая соответствует бездисперсионной модели, красная – дисперсионной

Заключение

Проведены тестовые расчеты распространения уединенной волны различной полярности в рамках теории мелкой воды и уравнений Буссинеска. Вычисления сделаны с помощью программного комплекса CLAWPACK-BOUSSCLAW. Полученные результаты соответствуют имеющимся теоретическим представлениям и позволяют судить о пригодности вычислительного комплекса CLAWPACK-BOUSSCLAW для решения практических задач.

Представленные результаты получены в рамках гранта РФФИ 20-05-00162.

Библиографический список

1. Стокер Дж. Волны на воде / Дж. Стокер. – М.: ИЛ, 1959. – 618 с.
2. Лайтхилл Дж. Волны в жидкостях / Дж. Лайтхилл. – М.: Мир, 1981. – 598 с.
3. Whitham G. B. Linear and Nonlinear Waves / G. B. Whitham. – N.Y.: Wiley-Interscience, 1974. – 636 p.
4. Tinti S. Analytical evolution of tsunamis induced by near-shore earthquakes on a constant-slope ocean / S. Tinti, R. Tonini // Journal of Fluid Mechanics. – 2005. – V. 535. – P. 33-64. DOI: 10.1017/S0022112005004532.
5. Родин А.А. Динамика длинных волн в прибрежной зоне моря с учетом эффектов обрушения / А.А. Родин, Е.Н. Пелиновский. – Н. Новгород: НГТУ, 2014. – 93 с.
6. Пелиновский Е.Н. Гидродинамика волн цунами / Е.Н. Пелиновский. – Н. Новгород: ИПФ РАН, 1996. – 276 с.
7. Khakimzyanov G. Dispersive Shallow Water Waves. Theory, Modeling, and Numerical Methods / G. Khakimzyanov, D. Dutykh, Z. Fedotova, O. Gusev. – Cham: Birkhäuser, 2020. – 284 p. DOI: 10.1007/978-3-030-46267-3.
8. Левин Б.В. Физика цунами и родственных явлений в океане / Б.В. Левин, М.А. Носов. – М.: Янус-К, 2005. – 360 с.
9. Kim J. A Boussinesq type extension of the GeoClaw model - a study of wave breaking phenomena applying dispersive long wave models / J. Kim, G.K. Pedersen, F. Løvholt, R.J. LeVeque // Coastal Engineering. – 2017. – V. 122. – P. 75-86. DOI: 10.1016/j.coastaleng.2017.01.005.
10. Madsen P.A. A new form of the Boussinesq equations with improved linear dispersion characteristics. Part 2. A slowly-varying bathymetry / P.A. Madsen, O.R. Sørensen // Coastal Engineering. – 1992. – V. 18, iss. 3-4. – P. 183-204. DOI: 10.1016/0378-3839(92)90019-Q.

11. Peregrine D.H. Long waves on a beach / D.H. Peregrine // Journal of Fluid Mechanics. – 1967. – V. 27, iss. 4. – P. 815–827. DOI: 10.1017/S0022112067002605.
12. LeVeque R.J. Finite Volume Methods for Hyperbolic Problems / R.J. LeVeque. – Cambridge: Cambridge University Press, 2002. – 558 p. DOI: 10.1017/CBO9780511791253.
13. Sommerfeld A. Partial differential equations in physics / A. Sommerfeld. – N.Y.: Academic Press inc., 1949. – 335 p.
14. Courant R. Supersonic Flow and Shock Waves / R. Courant, K.O. Friedrichs. – N.Y.: Interscience Publishers, 1999. – 464 p.

A.A. Rodin¹, N.A. Rodina², E.N. Pelinovsky^{1,3}

TEST CALCULATIONS OF THE PROPAGATION OF SURFACE WAVES IN THE FRAMEWORK OF THE NONLINEAR DISPERSION THEORY OF LONG WAVES

¹*Nizhny Novgorod State Technical University n.a. R.E. Alekseev,*

²*Lobachevsky State University of Nizhni Novgorod,*

³*Institute of Applied Physics RAS, Nizhny Novgorod*

Purpose: To investigate the effect of dispersion on the characteristics of the propagating wave in detail.

Design/methodology/approach: The calculations have been performed using the CLAWPACK-BOUSSCLAW software package which solves the two-dimensional Boussinesq equations.

Findings: Test computations of wave propagation on the surface of a liquid of finite depth are carried out within the framework of the nonlinear dispersion theory of long waves.

Originality/value: The results correspond to the available theoretical concepts and allow us to judge the suitability of the CLAWPACK-BOUSSCLAW computational complex for solving practical problems.

Keywords: long waves, numerical experiment, shallow water equations, dispersion, Boussinesq-type equations

**КОГЕРЕНТНЫЕ ВОЛНОВЫЕ СТРУКТУРЫ НА ПОВЕРХНОСТИ ВОДЫ,
ВОЗМОЖНОСТИ ИХ ОБНАРУЖЕНИЯ И «ВОЛНЫ-УБИЙЦЫ»**

¹ *Институт прикладной физики РАН,*

² *Национальный исследовательский университет "Высшая школа экономики",
г. Нижний Новгород*

Работа посвящена развитию способов прогноза волновых условий в море, в частности, описанию механизмов генерации аномально высоких волн («волн-убийц»). Важным объектом поверхностных волн являются долгоживущие нелинейные структуры, которые ранее считались практически нереализуемыми при реалистичных условиях в море. В работе показано, что такие структуры могут возникать при характерных для моря условиях случайных нелинейных волн, распространяться длительное время и приводить к возникновению очень больших волн. Показано, что эти структуры довольно хорошо соответствуют солитонам огибающей нелинейного уравнения Шредингера и могут быть найдены с помощью предложенного метода на основе решения прямой задачи рассеяния для этого уравнения.

Ключевые слова: нелинейные волны на поверхности воды, солитоны огибающей, численное моделирование, нелинейные волны на течении, экстремальные волны.

Введение

Интригующая проблема аномально высоких морских волн («волн-убийц») находится в центре внимания многочисленных исследований. Подход к изучению этого сложного природного явления с помощью прямого численного моделирования (Direct Numerical Simulation, DNS) нерегулярных волн стал популярным благодаря высокой производительности современных компьютеров и усовершенствованным численным алгоритмам. Потенциальные уравнения Эйлера хорошо подходят для моделирования создаваемых ветром гравитационных волн на поверхности моря. Они могут быть решены с помощью существующих численных алгоритмов во много раз быстрее, чем уравнения Навье-Стокса и спектральные уравнения Захарова, что позволяет моделировать ансамбли больших волн с учетом эффектов сильной нелинейности.

Наиболее интересными и пока практически неучитываемыми представляются нелинейные механизмы генерации «волн-убийц», связанные с образованием долгоживущих паттернов когерентных волн. Очень короткие долгоживущие группы в бесконечно глубокой воде с максимальной крутизной, близкой к порогу обрушения волн, были впервые обнаружены в численном моделировании исходных уравнений гидродинамики [1]. Такие сильно нелинейные гидродинамические солитоны огибающей были позднее воспроизведены в лабораторных условиях: сначала одиночные солитоны огибающей, а затем взаимодействующие пары солитонов огибающей [2, 3]. Однако эти солитонные группы распространялись по поверхности спокойной воды, не были окружены другими волнами.

Солитонные группы в полях нерегулярных волн

В наших предыдущих исследованиях проводилось стохастическое численное моделирование волн на поверхности воды со спектральными параметрами, типичными для океанских ветровых волн, с использованием псевдоспектрального метода высокого порядка (High Order Spectral Method, HOSM) [4, 5]. Эти численные эксперименты проводились с целью лучшего понимания механизмов генерации аномально больших волн и оценки характеристик «волн-убийц». По результатам серии численных экспериментов по моделированию однонаправленных волн со спектральными

параметрами, характерными для реального океана, были построены распределения вероятностей высот волн H , как показано на Рис. 1. Поскольку одним из важных вопросов является вопрос о достаточном размере статистического ансамбля, распределения вероятности были построены для разного числа реализаций: 100, 300 и 999. Каждая реализация представляла собой расчет эволюции цуга волн длиной примерно 10 км в течение 20 минут. Волны характеризовались спектром JONSWAP с умеренной пиковатостью $\gamma = 3$ и значительной высотой волн $H_{1/3} \approx 3.5$ м ($H_{1/3} \approx 4\sigma$, где σ – среднеквадратичное смещение поверхности η). Начальные условия соответствовали пиковому периоду $T_p = 10$ с; усредненный по ансамблю спектр волн в среднем слабо менялся за время моделирования.

Из Рис. 1 следует, что для ансамблей из 300 и 999 реализаций распределения вероятностей для отображенного интервала значений $H/H_{1/3}$ практически совпадают, в то время как для 100 реализаций функция заметно отличается в области $H/H_{1/3} > 2$. Отметим, что «волны-убийцы» обычно определяют с использованием именно такого порогового условия, $H/H_{1/3} \geq 2$. Пунктирной прямой на Рис. 1 построено распределение Рэлея, описывающее случай линейных волн с узким спектром. Видно, что распределение высот волн по данным численного моделирования совпадает с распределением Рэлея при $H/H_{1/3} > 2$ и заметно превышает распределение Рэлея в области больших высот.

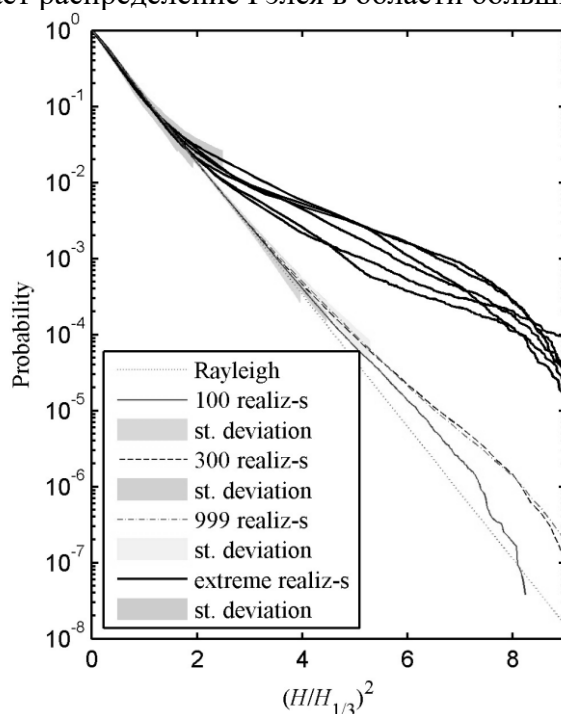


Рис. 1. Распределение вероятностей превышения высоты H по результатам численного моделирования однонаправленных волн. Тонкими линиями даны распределения, построенные по разным размерам статистических ансамблей; толстыми линиями – распределения для 6-ти самых экстремальных реализаций; пунктирная прямая линия соответствует распределению Рэлея.

Среди 999 реализаций было выделено 6, для которых наблюдалось трехкратное превышение значительной высоты, $H/H_{1/3} > 2$. Для них распределения вероятностей построены на Рис. 1 жирными линиями. Видно, насколько значительно эти реализации отличаются от среднего в смысле вероятностного распределения высот волн. Одна из этих реализаций, номер 295, была подробно исследована в работе [6], где было обнаружено, что из начальных условий рождается интенсивная волновая группа, которая распространяется без разрушения не только в течение 20 минут, но и вдвое дольше (более 240 периодов волн). Эта запись была затем проанализирована с помощью метода обратной задачи рассеяния в предположении локального приближения динамики волн

нелинейным уравнением Шредингера [7]. Возможность применения этого метода для солитонов огибающей большой крутизны была показана в работе [8]. Оказалось, что появившаяся на начальной стадии эволюции волн интенсивная группа может быть интерпретирована как солитон огибающей с амплитудой A_s , примерно равной значительной высоте волн, $A_s \approx 4\sigma$, см. Рис. 2а. Отметим, что на рисунке значения амплитуды выделенного солитона в соседние моменты времени примерно равны (что говорит об устойчивости процедуры обработки), а само значение амплитуды солитона плавно изменяется в небольшом диапазоне в течение всех 240 периодов волн. Некоторые (не все) из выделенных выше 6-ти реализаций также характеризуются возникновением интенсивных солитонов огибающей с амплитудой до порядка 4σ .

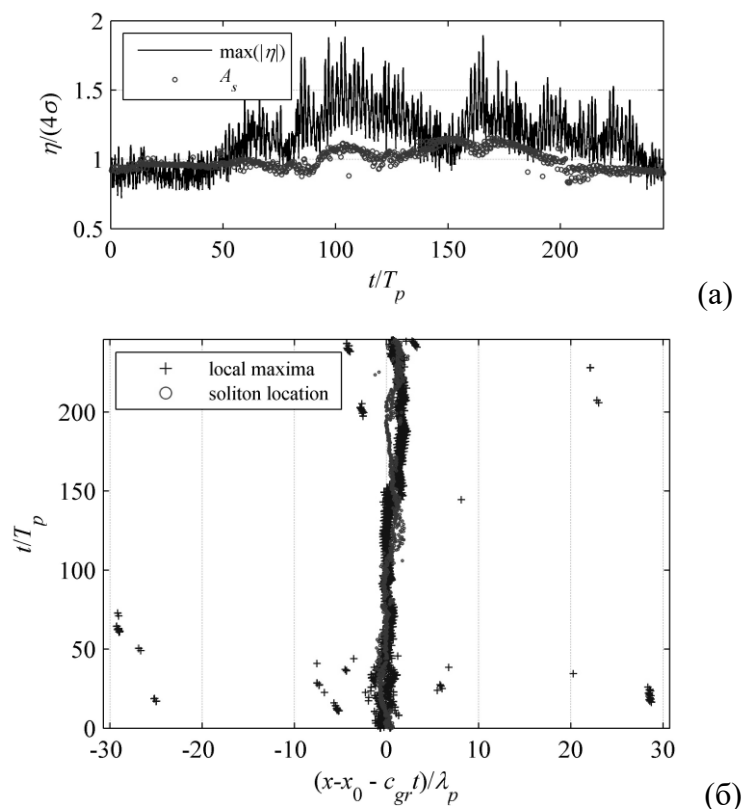


Рис. 2. Изменение во времени максимума смещения поверхности $\max|\eta|$ в расчетной области и амплитуды A_s максимальной солитонной группы для реализации 295 (а) и положение максимальной солитонной группы (кружки) и моментальных максимумов смещения поверхности (кресты) для той же реализации в сопровождающей системе координат (б).

Согласно Рис. 2а, первые 50 периодов максимальная высота гребней волн в области моделирования примерно равна амплитуде выделенного солитона, но позднее появляются смещения поверхности почти вдвое выше. На Рис. 2б разными символами построены положения выделенных солитонов огибающей (в сопровождающей системе отсчета, c_{gr} – групповая скорость волн, λ_p – длина волн, x_0 – некоторая константа) и положения моментальных максимумов смещения поверхности. Видно, что в абсолютном большинстве случаев экстремальные смещения поверхности происходят на фоне солитона огибающей. Наиболее высокий гребень в этом расчете соответствовал превышению $\max \eta / H_{1/3} \approx 1.9$, что является очень впечатляющей величиной.

Таким образом, выполненное численное моделирование и анализ его результатов демонстрируют возможность самопроизвольного формирования связанных нелинейных когерентных структур из нерегулярных однонаправленных волн на поверхности глубокой воды, которые могут существовать сотни периодов волн, приводят к экстремизации

волнения, а также представляют собой области, на фоне которых возникают аномально высокие волны. Эти когерентные структуры могут быть интерпретированы как короткие солитоны огибающей, состоящие из крутых волн (гидродинамические солитоны) и могут быть выделены в поле нерегулярных волн с помощью специального метода, использующего аппарат задачи рассеяния для нелинейного уравнения Шредингера [8].

Солитоны огибающей в полях захваченных волн

Нелинейное уравнение Шредингера (НУШ) для волн на глубокой воде может быть записан в форме

$$i\left(\frac{\partial A}{\partial t} + c_{gr} \frac{\partial A}{\partial x}\right) + \frac{\omega_0}{8k_0^2} \frac{\partial^2 A}{\partial x^2} + \frac{\omega_0 k_0^2}{2} A|A|^2 = 0, \quad (1)$$

где k_0 и ω_0 – волновое число и частота несущей волны, связанные дисперсионным соотношением $\omega_0^2 = gk_0$, g – ускорение свободного падения, а комплекснозначная функция $A(x, t)$ описывает поле волн. В частности, смещение в первом приближении записывается по формуле $\eta = \text{Re}(A \exp(i\omega_0 t - ik_0 x))$.

Для волн на глубокой воде НУШ (1) оказывается фокусирующего типа, что обеспечивает возможность модуляционной неустойчивости цугов интенсивных волн, образования солитонов огибающей и, в конечном итоге, повышения вероятности высоких волн в сравнении с гауссовой статистикой (эффект «волн-убийц»). Известно, что с учетом углового спектра волн модуляционная неустойчивость быстро ослабляется, так что для реалистичных (не однонаправленных) морских волн роль модуляционной неустойчивости в формировании высоких волн все еще остается под вопросом.

В то же время, в случае захвата волн особенностями распределения течений или батиметрии размерность системы эффективно снижается, даже если формально волны характеризуются не узким спектром направлений. Такая задача может быть описана в рамках модового подхода. В частности, в литературе было получено НУШ для краевых волн [9], а также для волн на встречных струйных течениях [10]. В последнем случае НУШ записывается в форме

$$i\left(\frac{\partial A}{\partial t} + V \frac{\partial A}{\partial x}\right) + \frac{\omega}{8k_0^2} \frac{\partial^2 A}{\partial x^2} + \kappa \frac{\omega k_0^2}{2} A|A|^2 = 0, \quad (2)$$

$$\kappa = \frac{\int_{-\infty}^{\infty} Y^4 dy}{\int_{-\infty}^{\infty} Y^2 dy}, \quad V = \frac{1}{\int_{-\infty}^{\infty} Y^2 dy} \int_{-\infty}^{\infty} \left(\frac{g^2 k_0}{2(\omega - k_0 U)^3} + U \right) Y^2 dy,$$

где $U(y)$ – профиль поперечного течения, которое предполагается однородным по глубине, а частота ω и поперечная мода $Y(y)$ находятся через решение соответствующей краевой задачи. Заметим, что уравнение (2) отличается от (1) только значениями коэффициентов ω и V (которые несильно отличаются от ω_0 и c_{gr} при вполне реалистичных предположениях), а также коэффициентом $0 < \kappa < 1$, который несколько снижает инкремент и область неустойчивых возмущений модуляционной неустойчивости, но не отменяет ее. Поскольку уравнение НУШ сохраняет свою форму, то могут быть использованы и классические результаты для интегрируемого НУШ, в частности, может быть явно записано солитонное решение для (2). Оказывается, что при вполне разумных предположениях на параметры течения краевая задача на ω и $Y(y)$ может быть решена точно для некоторых профилей $U(y)$ (или численно в общем случае). Тогда можно записать в аналитической или полуаналитической форме солитонные решения для трехмерной задачи волн на встречном струйном течении.

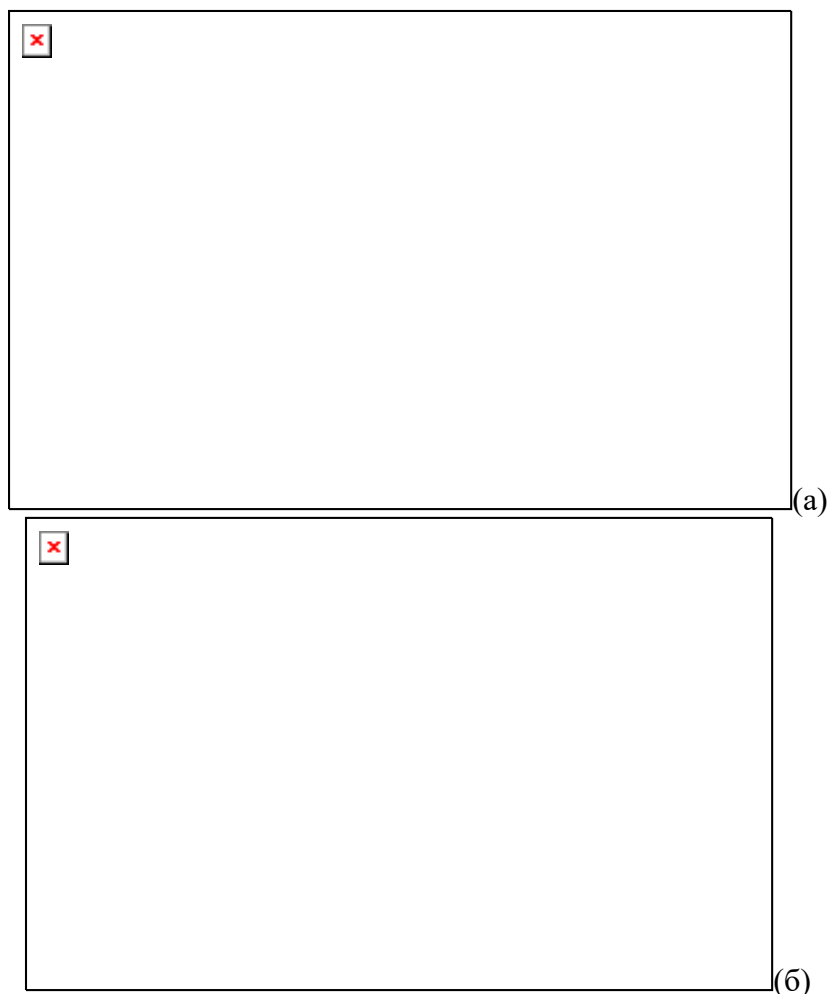


Рис. 3 Расчет эволюции короткого солитона огибающей с характерной крутизной $k_0 \max|A| \approx 0.24$ на встречном струйном течении: смещение поверхности в начале расчета (а) и после одного прохода периодической области (б). Профиль течения показан стрелками. По краям рисунков построены продольные и поперечные профили решений.

Несмотря на ряд использованных приближений и предположений, полученные решения не просто довольно хорошо описывают слабонелинейные волны, когда солитон огибающей представлен длинным цугом волн небольшой крутизны, но способны описывать решение уравнений гидродинамики даже в случае неожиданно крутых волн, как показано на Рис. 3. Смещение поверхности, соответствующее начальному условию, построено на Рис. 3а. Профиль течения, близкий к функции $\text{sech}^2(y/Lu)$, где Lu характерная полуширина течения, показан стрелками. Течение направлено навстречу оси Ox , вдоль которой распространяется солитон огибающей. Учитывая большую крутизну солитона, $k_0 \max|A| \approx 0.24$, он также соответствует и сильной дисперсии. Солитон локализован вдоль обоих горизонтальных направлений (x, y), состоит из всего нескольких осцилляций волн, а поперечный размер солитона примерно равен продольному. Эволюция солитона рассчитывалась в рамках потенциальных уравнений Эйлера (кодом HOSM) с учетом сильной нелинейности волн и полной дисперсии. Использовались периодические условия вдоль обеих горизонтальных осей. На Рис. 3б показано смещение поверхности после того, как солитон пробежал полный круг вдоль оси Ox . По краям расчетной области профили смещения поверхности в этот момент времени (сплошная линия) сопоставлены с профилями начального условия (точки). Видно, что солитон на струйном течении потерял часть своей энергии. Судя по графику поверхности на Рис. 3а, большая часть излученной энергии остается захваченной течением. При этом группа сохранила свою структуру, пройдя расстояние в 20 длин волн или несколько характерных длин группы.

Можно ожидать, что солитоны огибающей, возникающие на струйных течениях или на особенностях батиметрии, формирующих волновод, будут также являться местом возникновения аномально высоких волн, а потому возможность их выделения в поле нерегулярных волн, подобная описанной выше для однонаправленных волн в однородных условиях, представляет интерес с точки зрения возможности прогноза экстремальных волн.

Исследования поддержаны проектом РФФИ №21-55-15008.

Библиографический список

1. Dyachenko A.I., Zakharov V.E. On the formation of Freak Waves on the surface of deep water // Письма в ЖЭТФ. 2008. Т. 88. С. 356-359.
2. Slunyaev A., Clauss G.F., Klein M., Onorato M. Simulations and experiments of short intense envelope solitons of surface water waves // Phys. Fluids. 2013. V. 25. P. 067105.
3. Slunyaev A., Klein M., Clauss G.F. Laboratory and numerical study of intense envelope solitons of water waves: generation, reflection from a wall and collisions // Phys. Fluids. 2017. V. 29. P. 047103.
4. Sergeeva A. Slunyaev A. Rogue waves, rogue events and extreme wave kinematics in spatio-temporal fields of simulated sea states // Nat. Hazards Earth Syst. Sci. 2013. V. 13. P. 1759-1771.
5. Slunyaev A., Sergeeva A., Didenkulova I. Rogue events in spatiotemporal numerical simulations of unidirectional waves in basins of different depth // Nat. Haz. 2016. V. 84. P. 549-565.
6. Slunyaev A.V., Kokorina A.V. Soliton groups as the reason for extreme statistics of unidirectional sea waves // J. Ocean Eng. Marine Energy. 2017. V. 3. P. 395-408.
7. Slunyaev A.V. Persistence of hydrodynamic envelope solitons: detection and rogue wave occurrence // arXiv: 2012.06201.
8. Слюняев А.В. Анализ нелинейного спектра интенсивного морского волнения с целью прогноза экстремальных волн // Изв. ВУЗов. Радиофизика. 2018. Т. 61. С. 1-23.
9. Pelinovsky E., Polukhina O., Kurkin A. Rogue edge waves in the ocean // Eur. Phys. J. Special Topics. 2010. V. 185. P. 35-44.
10. Shrira V.I., Slunyaev A.V. Nonlinear dynamics of trapped waves on jet currents and rogue waves // Phys. Rev. E. 2014. V. 89. P. 041002.

A.V. Slunyaev

COHERENT WAVE STRUCTURES ON THE WATER SURFACE, POSSIBILITY TO DETECT AND ROGUE WAVES

*Institute of Applied Physics of the Russian Academy of Sciences, Nizhny Novgorod
National Research University Higher School of Economics, Nizhny Novgorod*

The work is devoted to the development of methods for better forecasting of the wave conditions in the sea, in particular, to the description of the mechanisms of generation of abnormally high waves ("rogue waves"). An important object of surface waves is long-lived nonlinear structures, which were previously considered practically unrealizable under realistic conditions at sea. It is shown in the work that such structures can appear under conditions of random nonlinear waves typical for seas, propagate for a long time and lead to the occurrence of very large waves. It is shown that these structures quite well correspond to the envelope solitons of the nonlinear Schrödinger equation and can be found using the proposed method based on solving the direct scattering problem for this equation.

Keywords: nonlinear waves on the water surface, envelope solitons, numerical simulation, nonlinear waves on currents, extreme waves.

ЧИСЛЕННОЕ МОДЕЛИРОВАНИЕ БРИЗЕРНОЙ ТУРБУЛЕНТНОСТИ НА ФОНЕ СЛУЧАЙНОГО ШУМА КВАЗИГАРМОНИЧЕСКИХ ВОЛН

¹Институт прикладной физики РАН, ²Национальный исследовательский университет "Высшая школа экономики", Нижний Новгород

Бризеры (осциллирующие волновые пакеты) наравне с солитонами являются наиболее энергонесущими волнами в различных физических средах (к ним можно отнести поверхностные и внутренние волны, оптические сети). Поэтому бризеры во многом определяют общую волновую динамику волновых полей. В данной работе проведено численное моделирование ансамблей бризеров (так называемая бризерная турбулентность) на фоне случайного шума квазигармонических волн. Изучены характеристики рассмотренных волновых полей, и определены ситуации, наиболее опасные с точки зрения появления аномально больших волн (волн-убийц).

Ключевые слова: бризеры, бризерная турбулентность, асимметрия, эксцесс, модифицированное уравнение Кортевега – де Вриза.

Введение

Бризеры (осциллирующие волновые пакеты) и солитоны (уединенные волны) являются неотъемлемой частью различных физических систем. В связи со способностью сохранять свою энергию долгое время, они могут значительно изменять характеристики волновых полей. Изучению динамики систем, содержащих большое число взаимодействующих солитонов (солитонная турбулентность или солитонный газ), посвящено значительное количество работ [1-12] в рамках уравнения Кортевега – де Вриза и нелинейного уравнения Шредингера. Статистика бризерных полей остается пока менее изученной. Данная работа посвящена исследованию динамики ансамблей бризеров на фоне случайного шума квазигармонических волн с помощью численного моделирования в рамках модифицированного уравнения Кортевега – де Вриза. Изучены характеристики рассмотренных волновых полей, и определены ситуации, наиболее опасные с точки зрения появления аномально больших волн (волн-убийц).

Математическая модель

Модифицированное уравнение Кортевега – де Вриза:

$$\frac{\partial u}{\partial t} + 6u^2 \frac{\partial u}{\partial x} + \frac{\partial^3 u}{\partial x^3} = 0. \quad (1)$$

Точное решение в виде бризера имеет вид:

$$u(x, t) = -4q \operatorname{sech}(R) \left[\frac{\cos(f) - \frac{q}{p} \sin(f) \tanh(R)}{1 + \left(\frac{q}{p}\right)^2 \sin^2(f) \operatorname{sech}^2(R)} \right], \quad (2)$$

где

$$R = 2qx + 8q(3p^2 - q^2)t + R_0, \quad (3)$$

$$f = 2px + 8p(p^2 - 3q^2)t + f_0, \quad (4)$$

p, q, R_0, f_0 - свободные параметры. Параметр p влияет на количество осцилляций в бризере, q определяет амплитуду бризера. В случае если $q \gg p$ бризер может быть представлен суперпозицией двух мКДВ солитонов разной полярности. Если $q \ll p$ бризер может быть представлен солитоном огибающей нелинейного уравнения Шредингера.

Уравнение (1) решается численно псевдо-спектральным методом.

Бризерная турбулентность на фоне случайного шума

Рассмотрим влияние случайного шума квазигармонических волн на характеристики ансамбля бризеров. В начальный момент времени зададим волновое поле линейной суперпозицией 20 бризеров (2), расположенных на значительном расстоянии друг от друга. Их параметры: $f_i = 0, q_i = 0, R_i \in [-15: -30: -585], p_i \in [0.1: 0.1: 2], i = 1: 20$. R_i и p_i заданы случайным образом. Далее, на данное бризерное поле, мы с помощью линейной суперпозиции накладываем шум, заданный следующими формулами:

$$u(x,0) = \sum_{j=1}^{\frac{N}{2}-1} \sqrt{2S(k_j)\Delta k} \cos(k_j x + \varphi_j), \quad (5)$$

где $k_j = j\Delta k, j = 1, \dots, N/2 - 1$ – волновые числа, $\Delta k = 0.0052, N = 2^{15}$. Размер расчетной области выбран как $L = 1200$. Фазы φ_j случайно распределены в интервале $[0, 2\pi)$. Спектр $S(k_j)$ имеет форму Гаусса:

$$S(k) = rQ \exp\left(-\frac{(k - k_0)^2}{2K^2}\right), \quad k > 0. \quad (6)$$

Заданы следующие значения параметров: $K=0.1; k_0=1; Q = \frac{1}{4K\sqrt{2\pi}}$.

Фиксированное бризерное поле и пример случайного синусоидального поля представлены на рис. 1.

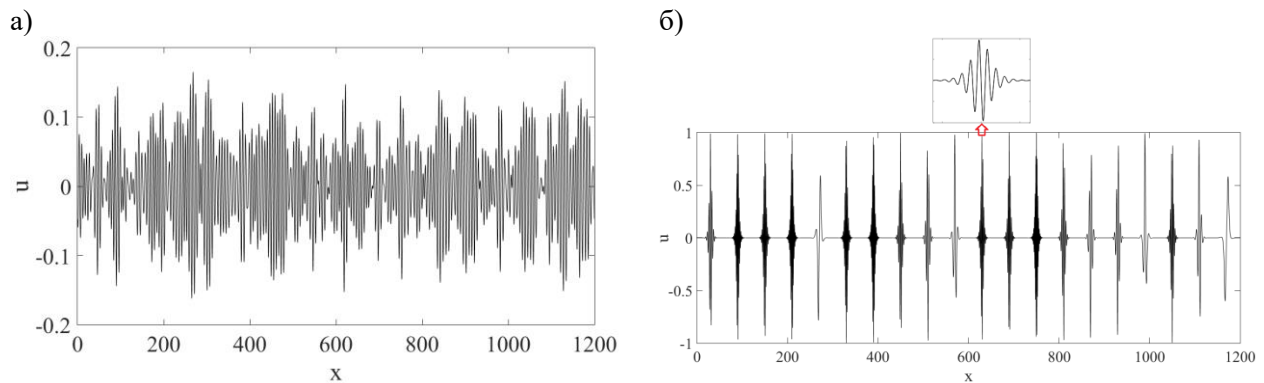


Рис. 1. Начальные волновые поля: а) синусоидальное поле ($r=0.05$), б) бризерное поле

Введем параметр $b = \frac{M_2 \text{ noise}}{M_2 \text{ br}}$, где $M_2 = \frac{1}{L} \int_{-\infty}^{\infty} u^2 dx$ – энергия каждой составляющей волнового поля (бризерной или шума). Рассмотрим несколько реализаций волновых полей с различными параметрами b : $b=0.1, 0.5, 1, 2, 5, 10$. Бризерный ансамбль фиксирован во всех шести случаях в момент времени $t=0$, меняется только уровень наложенного шума.

Квазигармонические волны сильно влияют на процесс эволюции бризеров, значительно увеличивая амплитуды волн (рис. 2).

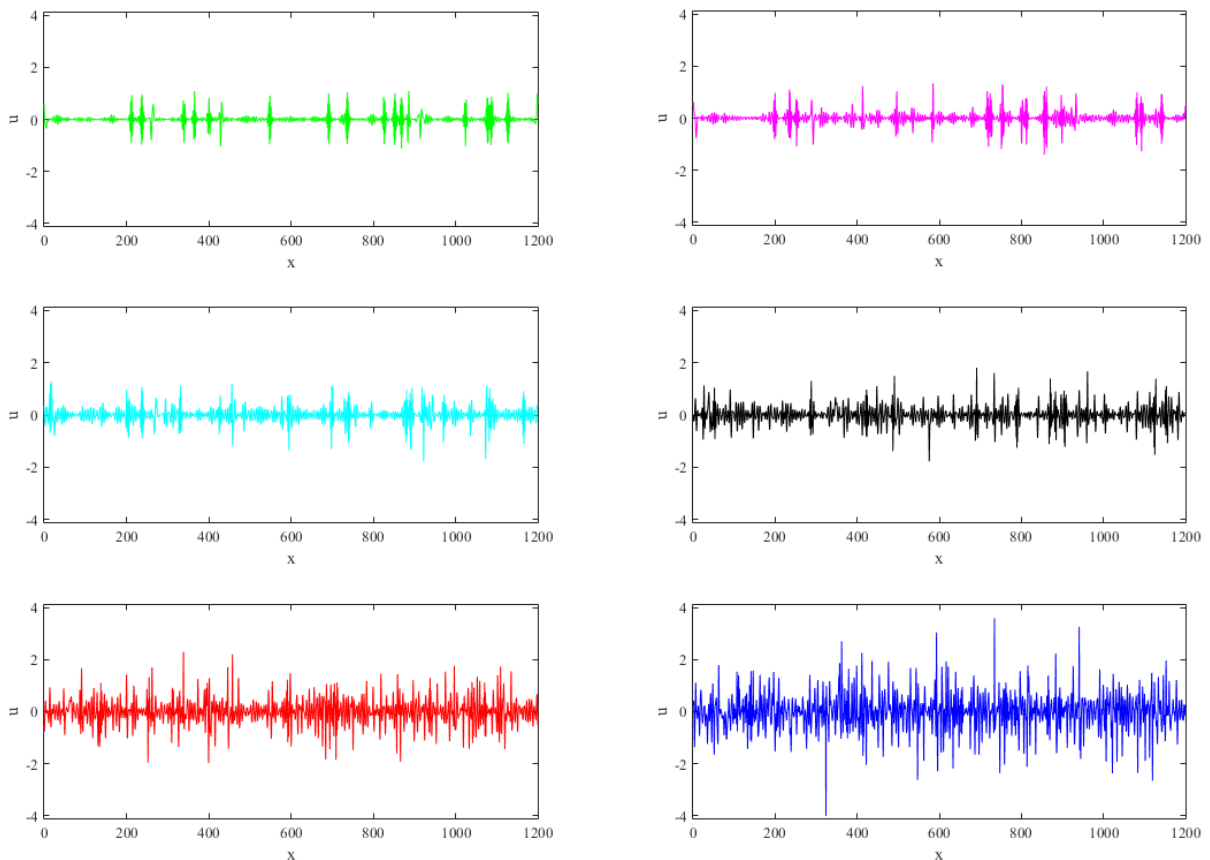


Рис. 2. Волновые поля при $t=100$ (зеленый - $b=0.1$, фиолетовый - $b=0.5$, голубой - $b=1$, черный - $b=2$, красный - $b=5$, синий - $b=10$)

Видно, что с увеличением параметра b , то есть с увеличением шума, амплитуды волн значительно возрастают. Если при $b=0.1$ среднее значение максимальных волн в каждый момент времени равняется 1.3, то при $b=10$ (когда энергия шума в 10 раз больше, чем энергия бризерной компоненты) это значение равняется 3.3 (рис. 3). Аналогичная ситуация с отрицательными значениями волновых полей. Причем рост усредненных экстремальных значений волновых полей происходит фактически по линейному закону.

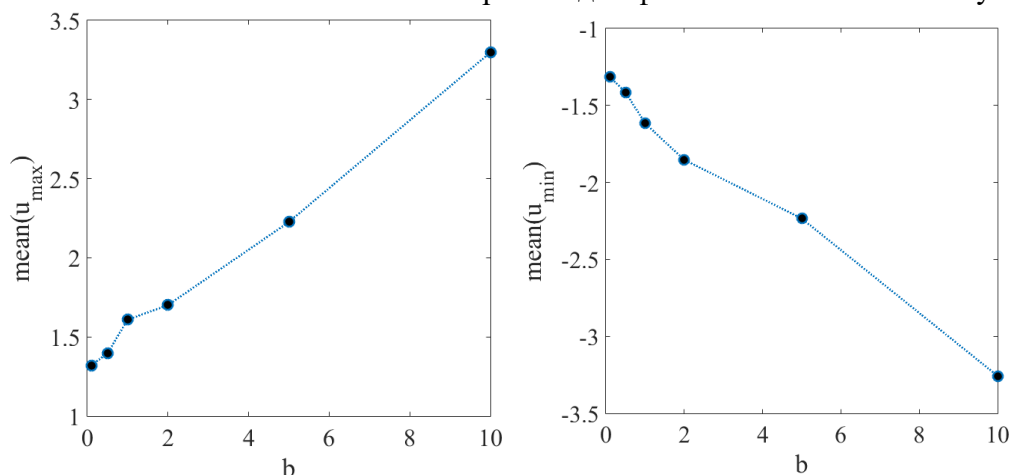


Рис. 3. Зависимость средних значений U_{\min} и U_{\max} от параметра b

Увеличение амплитуд и высот волн при увеличении шума также демонстрируется графиками функций распределения. Уже в начальный момент времени максимальные амплитуды (рис. 4) и высоты (рис. 5) волн в случае $b=10$ в два раза превосходят аналогичные параметры при $b=0.1$. А с увеличением времени, эта разница еще больше увеличивается: если сравнивать рис. 4а и 4б (или 5а и 5б) видно, что при $b=0.1...2$ функции распределения практически не меняются, в случае же $b=10$ (синяя кривая)

значения амплитуд и высот волн возрастают почти в полтора раза, что говорит о возникновении аномально больших импульсов в результате взаимодействия синусоидальных волн друг с другом и с бризерами.

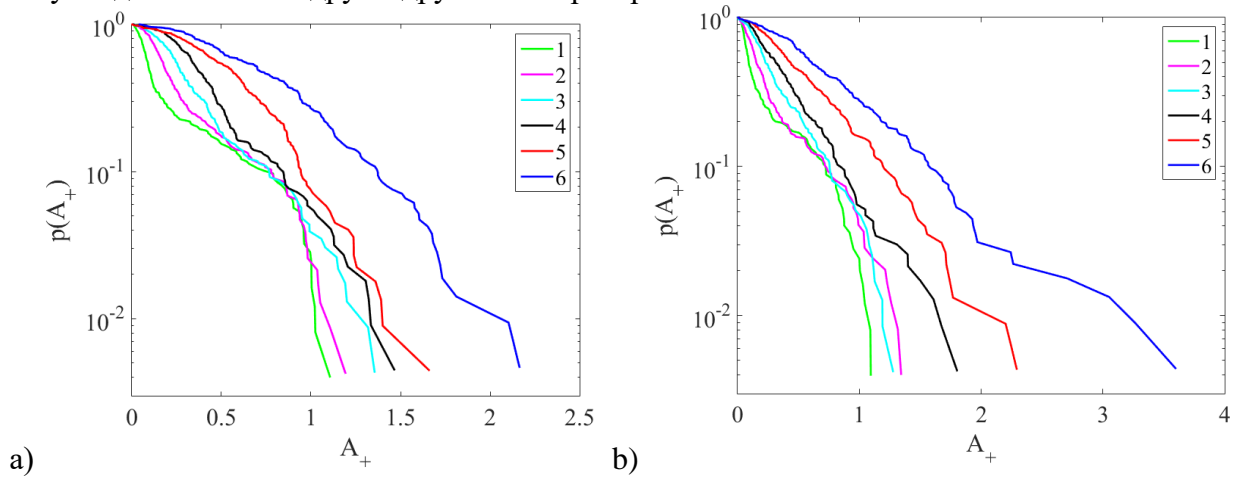


Рис. 4. Функции распределения положительных амплитуд волн: а) $t=0$, б) $t=100$

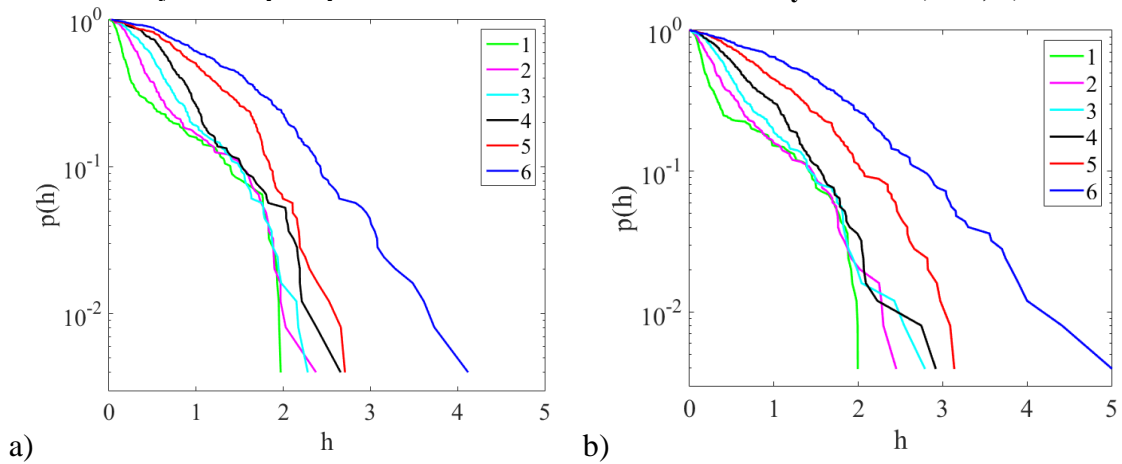


Рис. 5. Функции распределения высот волн: а) $t=0$, б) $t=100$.

Рассмотрим статистические характеристики волновых полей. В силу симметрии волновых полей третий момент (асимметрия) колеблется около нуля для всех волновых полей. Чем больше составляющая шума (с увеличением параметра b), тем ниже эксцесс, приближающийся к значению 3, соответствующему гауссовому процессу.

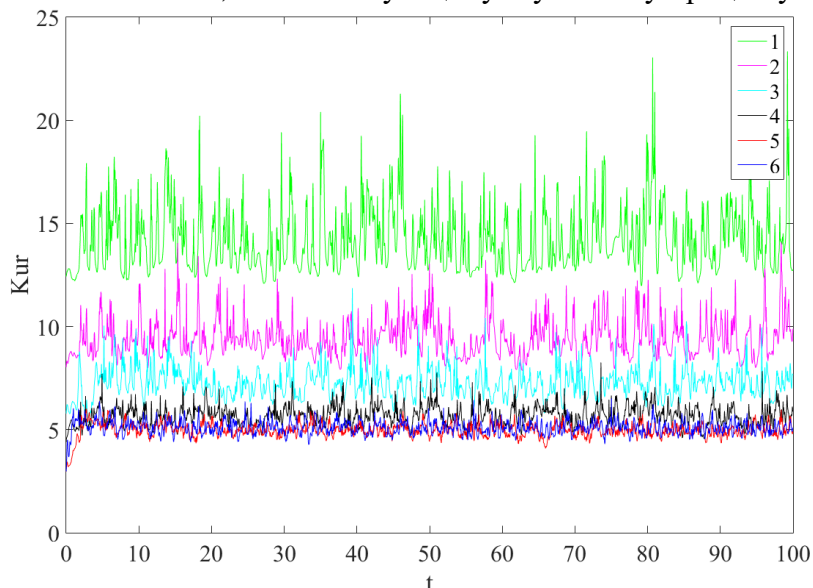


Рис. 6. Эволюция эксцесса волновых полей во времени

Заключение

Проведено численное моделирование динамики бризерного газа на фоне квазигармонических волн в рамках модифицированного уравнения Кортевега-де Вриза. Рассмотрено влияние шума различной интенсивности на характеристики волнового поля. Отмечается, что при увеличении энергии квазигармонических волн волновое поле становится более экстремальным: увеличиваются амплитуды и высоты волн. Экспесс же, наоборот, уменьшается с увеличением шума, так как волновое поле становится более гауссовым.

Работа выполнена при поддержке гранта РФФ 19-12-00253 «Нелинейные механизмы генерации волн-убийц».

Библиографический список

1. Захаров В.Е. Кинетическое уравнение для солитонов / В.Е. Захаров // ЖЭТФ. 1971. Т. 60. С. 993-1000.
2. El G.A. Critical density of a soliton gas. Chaos. 2016. V. 26. № 2. P. 1–7.
3. Agafontsev D.S. Integrable turbulence generated from modulational instability of cnoidal waves / D.S. Agafontsev, V.E. Zakharov // Nonlinearity. 2016. V. 29. № 11. P. 3551-3578.
4. Dudley J.M. Instabilities, breathers and rogue waves in optics / J.M. Dudley, F. Dias, M. Erkintalo, G. Genty // Nature Photonics. – 2014. – Vol. 8. – P. 755–764.
5. Akhmediev N. Breather turbulence versus soliton turbulence: Rogue waves, probability density functions, and spectral features / J.M. Soto-Crespo, N. Devine // Phys. Rev. E. 2016. V. 94. 022212.
6. Pelinovsky E. KDV soliton gas: interactions and turbulence, In: Challenges in Complexity: Dynamics, Patterns, Cognition (editors: I. Aronson, N. Rulkov, A. Pikovsky, L. Tsimring) / E. Shurgalina // Springer. 2017. P. 271-281.
7. Shurgalina E.G. Nonlinear dynamics of a soliton gas: Modified Korteweg-de Vries equation framework / E.G. Shurgalina, E.N. Pelinovsky // Phys. Lett. A. 2016. V. 380, № 24. P. 2049-2053.
8. Slunyaev A.V. Role of multiple soliton interactions in the generation of rogue waves: The modified korteweg-de vries framework / A.V. Slunyaev, E.N. Pelinovsky // Phys. Rev. Lett. 2016. V. 117. P. 214501.
9. Диденкулова Е.Г. Роль «толстого» солитона в динамике солитонного газа в рамках уравнения Гарднера / Е.Г. Диденкулова, Е.Н. Пелиновский // Известия Вузов. Радиофизика. 2018. Т. 61. С. 700–710.
10. Didenkulova (Shurgalina) E.G. Numerical modeling of soliton turbulence within the focusing Gardner equation: rogue wave emergence / E.G. Didenkulova (Shurgalina) // Physica D. 2019. V. 399. P. 35-41.
11. Шургалина Е.Г. Механизм образования волн-убийц в результате взаимодействия солитонов внутренних волн в стратифицированном водоеме / Е.Г. Шургалина // Известия РАН. Механика жидкости и газа. 2018. № 1. С. 61-67.
12. Slunyaev A.V. On the optimal focusing of solitons and breathers in long-wave models / A.V. Slunyaev // Studies in Applied Mathematics. 2019. V. 142. P. 385-413.

E.G. Didenkulova^{1,2}, O.I. Didenkulov¹

**NUMERICAL SIMULATION OF BREATER TURBULENCE ON THE BACKGROUND
OF RANDOM NOISE OF QUASI-HARMONIC WAVES**

¹⁾ Institute of Applied Physics RAS, ²⁾ HSE University, Nizhny Novgorod

Breathers (oscillating wave packets), along with solitons, are the most energy-carrying waves in various physical media (these include surface and internal waves, optical networks). Therefore, breathers largely determine the overall wave dynamics of wave fields. In this work, a numerical simulation of breather ensembles (the so-called breather turbulence) against the background of random noise of quasi-harmonic waves is carried out. The characteristics of the considered wave fields are studied, and the most dangerous situations from the point of view of the appearance of abnormally large waves (rogue waves) are determined.

Keywords: breathers, breather turbulence, skewness, kurtosis, modified Korteweg - de Vries equation

МОДЕЛИРОВАНИЕ ЦУНАМИ 1597 ГОДА НА РЕКЕ ВОЛГЕ С ПОМОЩЬЮ ОСРЕДНЕННЫХ ПО РЕЙНОЛЬДСУ УРАВНЕНИЙ НАВЬЕ-СТОКСА С УЧЕТОМ ОБРУШЕНИЯ

¹Федеральное государственное унитарное предприятие «Российский Федеральный Ядерный Центр Всероссийский научно-исследовательский институт экспериментальной физики» (ФГУП «РФЯЦ - ВНИИЭФ»)

²Нижегородский государственный технический университет им. Р. Е. Алексеева

В настоящее время для моделирования волн цунами главным образом применяются программы, основанные на теории мелкой воды. Система уравнений мелкой воды хорошо зарекомендовала себя при моделировании распространения волн цунами, однако она не способна воспроизвести сложную структуру трёхмерного течения, в частности, на стадии трансформации волны в шельфовой зоне при ее обрушении, накате на берег и продвижение по суше с взаимодействием с береговой инфраструктурой. Учесть особенности трехмерных структур цунами поможет использование численного моделирования, основанного на системе трехмерных уравнений Навье-Стокса, осредненных по Рейнольдсу.

В статье представлено описание метода расчета, основанного на осредненных по Рейнольдсу уравнений Навье-Стокса. Проведено численное моделирование оползневое цунами 1597 года на реке Волге с учетом обрушения волны на суше.

Ключевые слова: цунами, численное моделирование, осредненные по Рейнольдсу уравнения Навье-Стокса, метод VOF, пакет программ ЛОГОС

Введение

Одной из наиболее важных задач механики жидкости и газа является моделирование возникновения, распространения и процесса наката на берег волн цунами. Для оценки цунамиопасности разрабатываются методы моделирования волн цунами: их распространения вглубь суши и воздействия на прибрежную инфраструктуру. В настоящее время для моделирования волн цунами главным образом применяются программы, основанные на теории мелкой воды [1].

Система уравнений мелкой воды хорошо зарекомендовала себя при моделировании распространения волн цунами, однако она не способна воспроизвести сложную структуру трёхмерного течения. В частности, трехмерные структуры возникают на стадии трансформации волны в шельфовой зоне при ее обрушении, накате на берег и продвижение по суше с взаимодействием с береговой инфраструктурой.

Учесть особенности трехмерных структур цунами поможет использование численного моделирования, основанного на системе трехмерных уравнений Навье-Стокса, осредненных по Рейнольдсу [2,3]. Данная система является наиболее полной системой уравнений вязкой жидкости, учитывающей сложную структуру течений, возникающих в волнах цунами. Такая математическая модель позволяет единым образом моделировать движение и взаимное влияние твердой (оползень, тело), водной и воздушной сред.

Методика численного решения осредненных по Рейнольдсу уравнений Навье-Стокса для моделирования распространения поверхностных волн включает в себя метод Volume of Fluid (VOF) [4] для расчета течений со свободной поверхностью и набор «сжимающих» схем (HRIC, MCICSAM) [5,6] для сохранения минимально возможной толщины границы раздела сред. Данная методика широко используется для моделирования различных задач со свободной поверхностью, включая волны цунами.

Основные уравнения и метод расчета

Рассмотрим задачу о движении среды, состоящей из произвольного числа веществ с различными фазовыми состояниями. Примем во внимание упрощающие допущения о том, что все вещества описываются одним полем скорости и давления, фазовые переходы отсутствуют, рассматриваемые вещества являются несжимаемыми, отсутствуют стоки и источники, а также о том, что процессами теплообмена можно пренебречь. На данном этапе мы будем пренебрегать эффектами вращения Земли и ее сферичностью. В связи с этим система осредненных по Рейнольдсу уравнений Навье-Стокса в декартовых координатах примет вид (знаки осреднения опущены):

$$\begin{cases} \frac{\partial}{\partial x_i} (u_i) = 0 \\ \rho \frac{\partial u_i}{\partial t} + \rho \frac{\partial}{\partial x_j} (u_i u_j) = -\frac{\partial p}{\partial x_i} + \frac{\partial}{\partial x_j} (\tau_{ij} + \tau_{ij}^t) \\ \frac{\partial \alpha_w}{\partial t} + \frac{\partial}{\partial x_i} (u_i \alpha_w) = 0 \end{cases} \quad (1)$$

где i, j – нижние индексы, указывающие на принадлежность векторных компонент к декартовым координатам, $i, j = \{x, y, z\}$; ρ – плотность смеси, вычисляемая как $\rho = (\rho_w \alpha_w + \rho_a \alpha_a)$; w – (water) нижний индекс, указывающий на принадлежность к фазе «вода»; a – (air) нижний индекс, указывающий на принадлежность к фазе «воздух»; α_w – объёмная доля воды;

u_i – компонента вектора скорости, $i = \{x, y, z\}$; t – время; p – давление; x_i – компонента вектора декартовых координат, $i = \{x, y, z\}$; τ_{ij} – тензор вязких напряжений, который, согласно гипотезе Буссинеска, принимает вид:

$$\tau_{ij} = \mu \left(\frac{\partial u_i}{\partial x_j} + \frac{\partial u_j}{\partial x_i} + \frac{2}{3} \frac{\partial u_k}{\partial x_k} \delta_{ij} \right),$$

где μ – динамическая вязкость; δ_{ij} – дельта Кронекера; g_i – компонента вектора ускорения свободного падения.

Первые два уравнения системы (1) представляют собой уравнения сохранения массы и импульса. Третье уравнение – уравнение переноса объёмной доли жидкой фазы. Для фазы «воздух» уравнение переноса объёмной доли решать не требуется, поскольку, согласно принципам метода VOF, объёмная доля α_a вычисляется из соотношения $(\alpha_w + \alpha_a) = 1$.

Для моделирования оползневых цунами в системе уравнений (1) добавляется уравнение переноса объёмной доли оползневой фазы:

$$\frac{\partial \alpha_l}{\partial t} + \frac{\partial}{\partial x_i} (u_i \alpha_l) = 0, \quad (2)$$

а объёмная доля воздуха α_a вычисляется из соотношения $(\alpha_w + \alpha_l + \alpha_a) = 1$.

Данная система уравнений позволяет моделировать волн цунами, их распространение и последующий накат на берег, а также движение оползня с учетом реологии при наличии уравнения переноса объёмной доли для оползневой фазы. Использование лежащей в основе системы (1) методики VOF позволяет проводить численные эксперименты на сетках произвольной геометрической конфигурации.

Система уравнений (1) должна быть дополнена граничными условиями. Для задач цунами, как правило, используются граничные условия типа «стенка» для поверхности дна, неотражающие граничные условия на внешних границах акваторий. При моделировании наката никаких дополнительных граничных условий не требуется, используется только граничное условие «стенка» на подстилающей поверхности. На твердых стенках градиент давления и объёмных долей равен нулю:

$$\frac{\partial p}{\partial n} = 0, \frac{\partial \alpha_k}{\partial n} = 0, \quad (3)$$

значение скорости равно нулю:

$$u = 0, v = 0, w = 0, \quad (4)$$

т.е. невозможно ни проскальзывание жидкости вдоль границы «жидкость – твердая стенка», ни движение по нормали к ней. На границе «жидкость - жидкость» должны быть непрерывны скорость и сдвиговые напряжения. На верхней границе воздуха фиксируется нулевое статическое давление, градиенты скорости и объемных долей равны нулю:

$$\frac{\partial u}{\partial n} = 0, \frac{\partial v}{\partial n} = 0, \frac{\partial w}{\partial n} = 0, \frac{\partial \alpha_k}{\partial n} = 0, \quad (5)$$

Неотражающие граничные условия, необходимые для свободного ухода волн из расчетной области, будут подробно описаны в настоящей диссертации ниже. Начальные условия (моделирование очага цунами) будут представлены в последующих параграфах этой главы.

Дискретизация системы уравнений (1) осуществляется методом конечных объемов на произвольной неструктурированной сетке, а для её численного решения используется полностью неявный метод [7,8], основанный на известном алгоритме SIMPLE. Моделирование течений со свободной поверхностью подразумевает определенные модификации алгоритма SIMPLE. Описание основных формул модифицированного алгоритма SIMPLE, граничных условий и реализации в пакете программ ЛОГОС детально описано в [8,9]. Для ускорения расчетов распространения волн цунами на большие расстояния разработанная методика использует современные технологии ускорения расчетов и обладает эффективной масштабируемостью при использовании большого вычислительного поля процессоров. Это возможно благодаря применению многосеточных технологий [11], широко используемых в настоящее время в промышленных суперкомпьютерных вычислениях [12].

Моделирование оползневой цунами на реке Волге с учетом обрушения на берег

В 1597 г. на реке Волге в районе Нижнего Новгорода из-за подъема грунтовых вод произошел сход оползня, в результате которого возникла волна цунами [13]. Размеры сошедшего оползня предположительно составляли 300 на 200 м, объем не превышал 150000 м³, а его вес не превышал 410000 т (при плотности глинистых почв в 2700 кг/м³).

Для моделирования схода оползня и распространения образовавшихся волн методом отсечения была построена сеточная модель с количеством ячеек ~5 млн. Для более точного моделирования распространения волн сетка имеет сгущение к поверхности раздела фаз. Для детализации характера движения оползня был построен дополнительный блок измельчения в области его схода.

На рис. 1 показан процесс схода оползня в воду и образование волны. Максимальная высота волны в результате схода оползня составила почти 9 м (см. рис. 4)

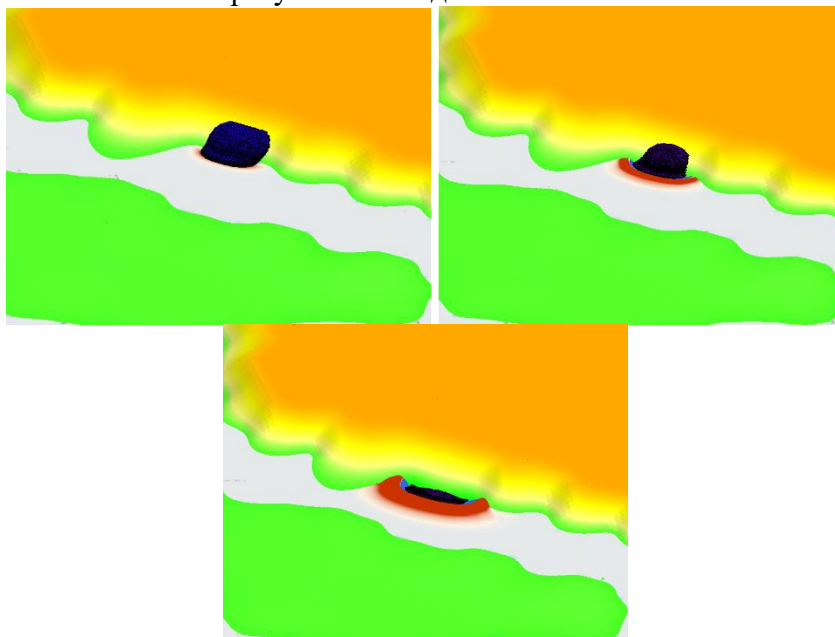


Рис. 1. Процесс схода оползня и образование волны

На рис. 2 представлены картины распространения цунами по реке на различные моменты времени. В [13] волна сначала зажата между высоким берегом Волги, откуда сошел оползень, и островом Печерские пески, затем волна выходит из узкой зоны и начинает распространяться как вниз, так и вверх по течению.

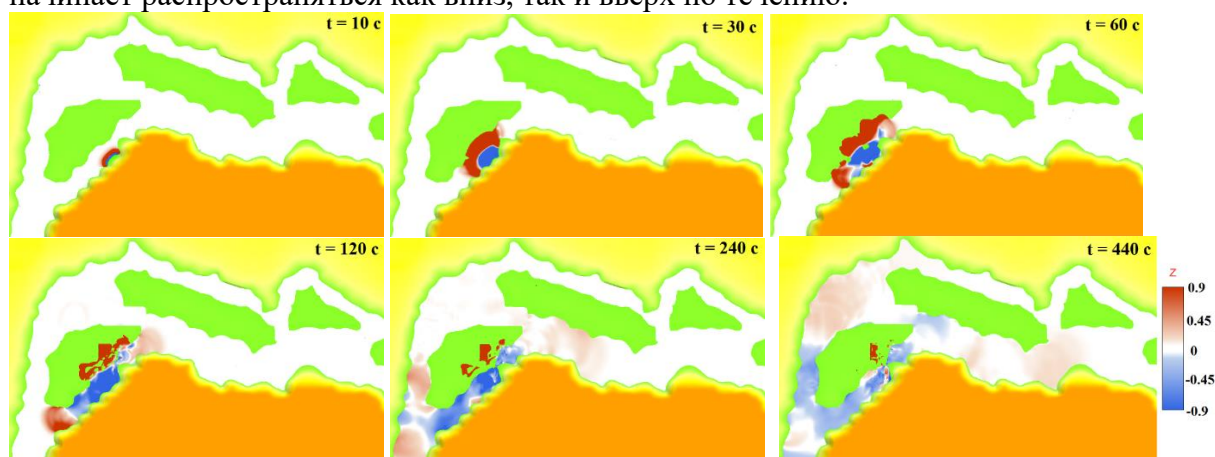


Рис. 2. Картины распространения цунами по реке на различные моменты времени

При расчете в рамках осредненных по Рейнольдсу уравнений Навье-Стокса основная волна распространяется в направлении схода оползня и набегает на противоположный берег. Затем волна также выходит из узкого русла реки и распространяется как вниз так и вверх по течению.

Из представленных рисунков видно, что после того, как волна дошла до берега острова, она начинает распространяться по суше вглубь острова на расстояние до 300 м.

Распространение волны по руслу реки описывается одинаково по всем моделям – модели, основанной на уравнениях мелкой воды, уравнениях Буссинеска и уравнениях Навье-Стокса. Понятно, что волна обрушается при приближении ее к берегу острова и накатывается на остров, после чего распространяется вглубь суши на рис. 3 приведены картины процесса распространения волны и ее обрушения. Рассматривалась область напротив схода оползня. Возможность воспроизвести процесс обрушения волны на берег и ее продвижение вглубь суши характеризует модель, основанную на осредненных по Рейнольдсу уравнениях Навье-Стокса как наиболее полную модель для моделирования всех стадий цунами.

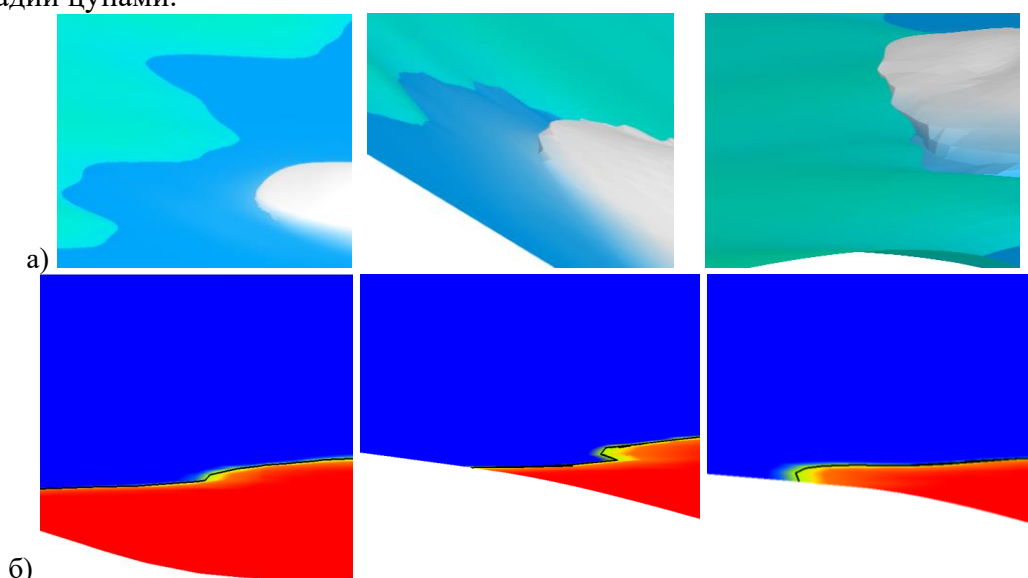


Рис. 3. Процесс распространения волны и ее обрушение на последовательные моменты времени (15 с, 20 с и 24 с): а) положение свободной поверхности, б) в сечении (черной линией отмечена изолиния по водной поверхности)

Рис. 3 показывает, что сначала волна распространяется по реке и обрушения не наблюдается, затем при подходе к берегу и уменьшении глубины реки начинается процесс обрушения и происходит накат на берег с обрушением волны.

На рис. 4 представлено распределения максимальных высот волн, полученные в рамках уравнений Навье-Стокса и уравнений мелкой воды. В обоих расчетах максимальная амплитуда волны наблюдается в месте схода оползня, однако при моделировании в рамках осредненных по Рейнольдсу уравнений Навье-Стокса основная волна распространяется в направлении схода оползня и набегают на противоположный берег, а не только вверх и вниз по течению реки.

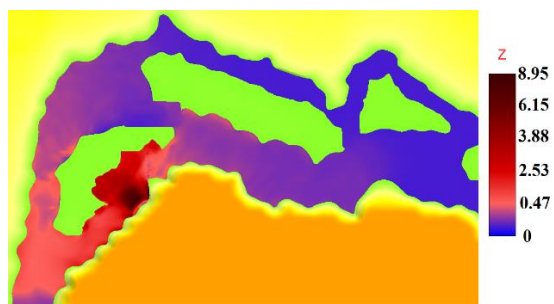


Рис. 4. Распределение максимальных высот

Таким образом, проведено сравнение результатов моделирования оползневых цунами на реке Волге, проведенное по двум пакетам программ – ЛОГОС (трехмерные уравнения Навье-Стокса) и TUNAMI (двумерные уравнения мелкой воды). Качественно распространение волн по обоим методам описывается идентично. Однако при использовании метода, основанного на осредненных по Рейнольдсу уравнениях Навье-Стокса, есть возможность учесть обрушение волн при возбуждении волн сходом оползня, их распространении по реке и накате на берег.

Работа выполнена при финансовой поддержке гранта Президента Российской Федерации по государственной поддержке научных исследований ведущих научных школ Российской Федерации НШ-2485.2020.5, а также при финансовой поддержке гранта РФФИ 20-05-00162.

Библиографический список

1. Goto S., Ogawa Y., Shuto N., Imamura N. Numerical method of tsunami simulation with the leap-frog scheme (IUGG/IOC Time Project), IOC Manual, UNESCO, 1997, № 35, P. 96 p.
2. Ландау Л.Д., Лифшиц В.М. Гидродинамика. – М.: Наука, 1988.
3. Лойцянский Л.Г. Механика жидкости и газа. – М.: Наука, 1973.
4. Hirt C.W., Nichols B.D. Volume of Fluid (VOF) method for the dynamics of free boundaries // Journal of Computational Physics, 1981, V. 39, P. 201-225.
5. Muzaferija S., Peric M., Sames P., Schelin T. A twofluid Navier-Stokes solver to simulate water entry // 22 Symp. on Naval Hydrodynamics, 1998.
6. Waclawczyk T., Koronowicz T. Remarks on prediction of wave drag using VOF method with interface capturing approach // Archives of civil and mechanical engineering, 2008, V.8, P.5-14.
7. Chen Z.J., Przekwas A.J. A coupled pressure-based computational method for incompressible/compressible flows // Journal of Computational Physics, 2010, V. 229, P.9150-9165.
8. Козелков А.С., Мелешкина Д.П., Куркин А.А., Тарасова Н.В., Лашкин С.В., Курулин В.В. Полностью неявный метод решения уравнений Навье-Стокса для расчета многофазных течений со свободной поверхностью // Вычислительные технологии, 2016, Т. 21, №5, С. 54-76.

9. Козелков А.С. Методика численного моделирования цунами оползневого типа на основе уравнений Навье-Стокса // Вычислительная механика сплошных сред, 2016, Т. 9, №2, С. 218-236.
10. Moukalled F., Darwish M. Pressure-Based Algorithms for Multi-Fluid Flow at All Speeds- Part I: Mass Conservation Formulation, Numerical Heat Transfer, Part B: Fundamentals, 2004, V. 45, P. 495-522.
11. Волков К.Н., Дерюгин Ю.Н., Емельянов В.Н., Карпенко А.Г., Козелков А.С., Тетерина И.В. Методы ускорения газодинамических расчетов на неструктурированных сетках. – Москва: Физматлит, 2013, 536 с.
12. Козелков А.С., Шагалиев Р.М., Курулин В.В., Ялозо А.В., Лашкин С.В. Исследование потенциала суперкомпьютеров для масштабируемого численного моделирования задач гидродинамики в промышленных приложениях // Вычислительная математика и математическая физика, 2016, Т. 56, № 8, С. 1524–1535.
13. Диденкулова И.И., Зайцев А.И., Красильщиков А.А., Куркин А.А., Пелиновский Е.Н., Ялчинер А.С. Нижегородское цунами 1597 г. На реке Волге // Известия АИИ им. А.М. Прохорова. Прикладная математика и механика. 2003. Т. 4. С. 170-184.

E.S. Tyatyushkina^{1,2}

SIMULATIONS OF THE 1597 TSUNAMI WAVE AND ITS RIVERSIDE COLLAPSE ON THE VOLGA RIVER USING THE REYNOLDS-AVERAGED NAVIER-STOKES EQUATIONS

¹*Federal State Unitary Enterprise “Russian Federal Nuclear Center – All-Russian Research Institute of Experimental Physics” (FSUE “RFNC-VNIIEF”)*

²*Nizhny Novgorod State Technical University n.a. R.E. Alexeev*

At present, simulations of tsunami waves are mostly performed by codes based on the shallow water theory. The system of shallow-water equations has demonstrated its efficiency in simulations of tsunami wave propagation, but it is not as good at reproducing the complex three-dimensional flow structure, in particular, transformation of tsunami waves in the shelf sea area during their collapse, run-up and inland movement, including their interaction with the coastal infrastructure. Numerical simulations based of the system of three-dimensional Reynolds-averaged Navier-Stokes equations can help resolve the details of such three-dimensional tsunami structures.

The paper describes a computational technique employing the Reynolds-averaged Navier-Stokes equations and presents numerical simulations of the 1597 landslide tsunami wave on the Volga River, including its riverside collapse.

This work was funded by grants of the President of the Russian Federation for state support of research projects by leading scientific schools of the Russian Federation (NSh-2485.2020.5) and by the Russian Foundation for Basic Research (RFFI Project 20-05-00162).

Keywords: tsunami, numerical simulation, Reynolds-averaged Navier-Stokes equations, VOF method, LOGOS software package

РЕКОНСТРУКЦИЯ СОБЫТИЙ «ВОЛН-УБИЙЦ» У БЕРЕГОВ О. САХАЛИН¹ *Институт прикладной физики РАН,*² *Национальный исследовательский университет "Высшая школа экономики",**г. Нижний Новгород*

Настоящая работа посвящена анализу волн-убийц, зарегистрированных у берегов о. Сахалин, и реконструкции этих событий в рамках прямого численного моделирования. На основе данных долгосрочных измерений смещения поверхности выделены несколько временных записей, содержащих аномально высокие волны, которые более чем в 2 раза превышают характерную высоту волнения, и численно промоделирована пространственная эволюция таких событий в рамках слабонелинейной модели модулированных волн с учетом переменной батиметрии. Оценены геометрические параметры экстремальных волн и длительности их существования.

Ключевые слова: численное моделирование, нелинейные волны, натурные данные, экстремальные волны, статистика.

Введение

Измерения высот ветровых волн на поверхности моря чаще всего производят в одной точке, такие временные записи дают ограниченную информацию о волнах, поскольку не отражают их динамику. Используя предположение об однонаправленности поверхностных волн, их динамику можно рассматривать как одномерную, и распространение таких волн может быть реконструировано. В качестве приближенной модели могут быть использованы эволюционные уравнения для слабонелинейных слабомодулированных волн. В результате будет получена не только более подробная информация о волнах вблизи точки измерения, но также восстановлена их эволюция на некотором удалении от точки регистрации. В настоящей работе исследуются записи, полученные в Охотском море вблизи от берега острова Сахалин, когда изменением глубины моря пренебречь нельзя. Поэтому для моделирования мы используем модификацию эволюционного уравнения, учитывающую эффект переменной батиметрии. В работе рассмотрено три временные записи поверхностных волн, содержащие волны-убийцы согласно критерию превышения высоты волн. Чтобы понять нелинейную динамику экстремальных волн, распространяющихся со стороны открытого моря к берегу, и оценить опасность волн-убийц на побережье, мы реконструируем эволюцию таких экстремальных событий с использованием нелинейного уравнения Шредингера с переменными коэффициентами.

Натурные измерения

В течение нескольких лет Специальным конструкторским бюро средств автоматизации морских исследований ДВО РАН проводились непрерывные измерения колебаний уровня моря у берегов о. Сахалин в районе мыса Свободный. Измерения осуществлялись автономными донными регистраторами гидростатического давления. Датчики были установлены на глубине около 8 м. Пересчет вариаций давления в вертикальное смещение морской поверхности производился по линейной теории. Данные оказались очень неоднородными, что определялось метеорологическими условиями в данном районе (силой и направлением ветра, полем атмосферного давления). Все натурные записи делились на отрезки по 20 мин, и предполагалось, что колебание смещений поверхности внутри 20-минутного интервала является стационарным случайным процессом.

Для настоящего исследования были отобраны три 20-ти минутные реализации, содержащие anomalно высокие волны, обозначенные на рис. 1 как А, В и С. Для определения волн-убийц использовался общепринятый критерий:

$$AI \equiv \frac{H}{H_s} > 2, \quad (1)$$

где H – высота одиночной волны, а H_s – значительная высота волны. На графиках на рис. 1 справа построены выделенные anomalно высокие волны, демонстрирующие разные геометрические формы: так называемая «дыра в море» с характерной глубокой впадиной, волна знакопеременной формы и волна с высоким гребнем. Измерения проводились на расстоянии 500 м от берега на глубине 8 м. Характерные периоды волн $T_p = 8$ с, а соответствующая безразмерная глубина $k_p h = 0.76$, где k_p определяется через дисперсионное соотношение для гравитационных волн на конечной глубине $\omega^2 = gk \tanh(kh)$. Значительные высоты волн, высоты волн-убийц и высоты гребней H_{fr}^{cr} приведены в Таблице 1. Характерная крутизна волны, оцененная по среднеквадратическому отклонению η_{rms} смещения поверхности как $\varepsilon = k_p \eta_{rms}$, изменяется в диапазоне от 0.01 до 0.02, что соответствует относительно слабой нелинейности. Крутизна выявленных волн-убийц $k_{fr} H_{fr}^{cr}$ составляет от 0.07 до 0.1. Таким образом, идентифицированные волны-убийцы также характеризовались слабой нелинейностью.

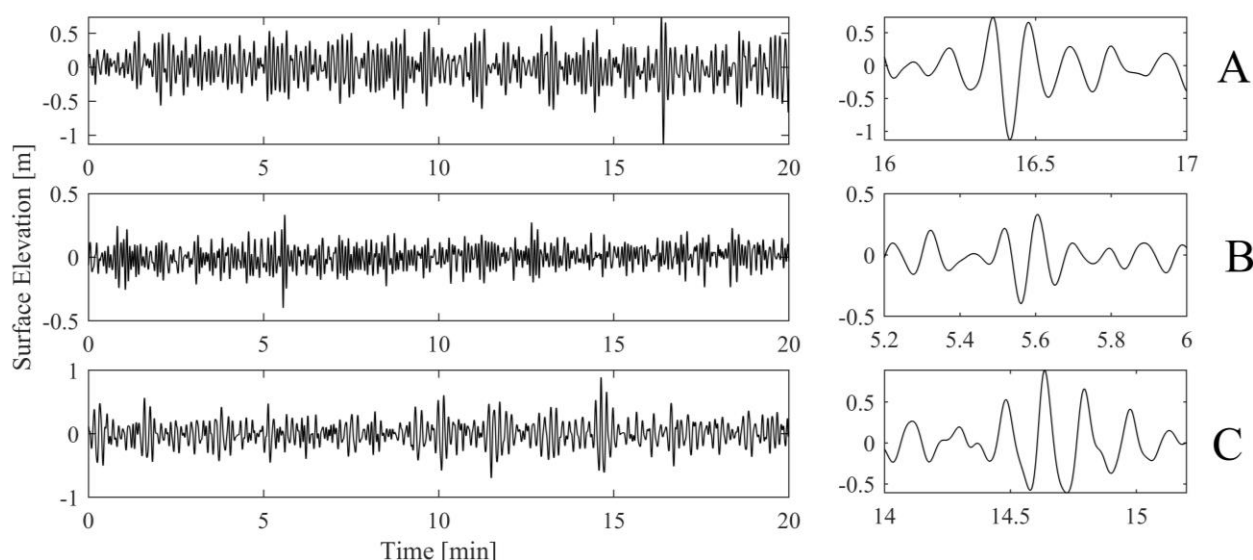


Рис. 1. Натурные 20-ти минутные записи смещения поверхности, содержащие anomalно высокие волны.

Для численного моделирования мы использовали реальную батиметрию от точки измерения к суше, интерполированную с использованием полиномиальной функции, а распределение глубин от точки измерения в сторону моря считали изменяющимся по линейному закону.

Таблица 1.

Основные параметры натурных измерений для серий А, В, С.

	ε	$k_p h$	T_p [s]	T_{fr} [s]	λ_{fr} [m]	H_{fr} [m]	H_{fr}^{cr} [m]	H_s [m]	Ur
А	0.02	0.76	8.1	7.4	59	1.9	0.74	0.9	9.1
В	0.01	1.19	5.7	5.5	40	0.73	0.33	0.33	1.63
С	0.015	0.62	9.7	9.3	78	1.6	0.89	0.75	10.7

Реконструкция волн-убийц в рамках уравнения Шредингера с переменными коэффициентами

Для восстановления эволюции волн, распространяющихся в одном направлении (в сторону берега), в окрестности точки измерения мы используем нелинейное уравнение Шредингера (НУШ) с учетом переменной батиметрии. Влияние переменного дна описывается с помощью коэффициентов, зависящих от глубины [1]

$$i \frac{dB}{dx} = -i\mu \frac{d(kh)}{dx} B - \frac{i}{c_g} \frac{dB}{dt} + \lambda \frac{d^2 B}{dt^2} + \nu |B|^2 B, \quad (2)$$

$$\mu = \frac{(1-\sigma^2)(1-kh\sigma)}{\sigma + kh(1-\sigma^2)}, \quad \lambda = \frac{1}{2c_g \omega_0} \left[1 - \frac{gh}{c_g^2} (1-kh\sigma)(1-\sigma^2) \right],$$

$$\nu = \frac{\omega_0 k^2}{16\sigma^2 c_g} \left[9 - 10\sigma^2 + 9\sigma^4 - \frac{2\sigma^2 c_g^2}{(gh - c_g^2)} \left(4 \frac{c_p^2}{c_g^2} + 4 \frac{c_p^2}{c_g^2} (1-\sigma^2) + \frac{gh}{c_g^2} (1-\sigma^2)^2 \right) \right],$$

$$\sigma = \tanh(kh), \quad c_p = \frac{\omega_0}{k}, \quad c_g = \left. \frac{\partial \omega}{\partial k} \right|_{\omega=\omega_0} = \frac{1}{\sqrt{4\omega_0}} \left(g\sigma + \frac{ghk}{ch^2(kh)} \right).$$

Здесь $B(x, t)$ – комплексная огибающая, x – горизонтальная координата, направленная в сторону берега, $x = 0$ соответствует береговой линии, а измерительный датчик расположен в точке $x = x_0 > 0$, t – время, g – ускорение свободного падения, c_p и c_g – локальная фазовая и групповая скорости соответственно. Когда глубина воды h постоянна, уравнение (2) сводится к классическому уравнению НУШ для волн в бассейне постоянной глубины. Предполагается, что волны распространяются строго к берегу, несущая частота ω_0 остается постоянной, а волновое число k изменяется в соответствии с дисперсионным соотношением для волн на конечной глубине.

Уравнение (2) решается численно с использованием модификации метода сплит-степ-Фурье, аналогичный алгоритм использовался в нашей работе [2]. Эволюция волны в обратном направлении (от берега) моделируется с использованием отрицательного шага интегрирования в пространстве. В качестве начальных условий используются натурные временные записи смещения волновой поверхности в точке $x = x_0$. Несущая частота устанавливается равной средней частоте энергетического волнового спектра. Комплексная огибающая $B(t)$ при $x = x_0$ ищется с помощью итерационной процедуры, подробно описанной в [3].

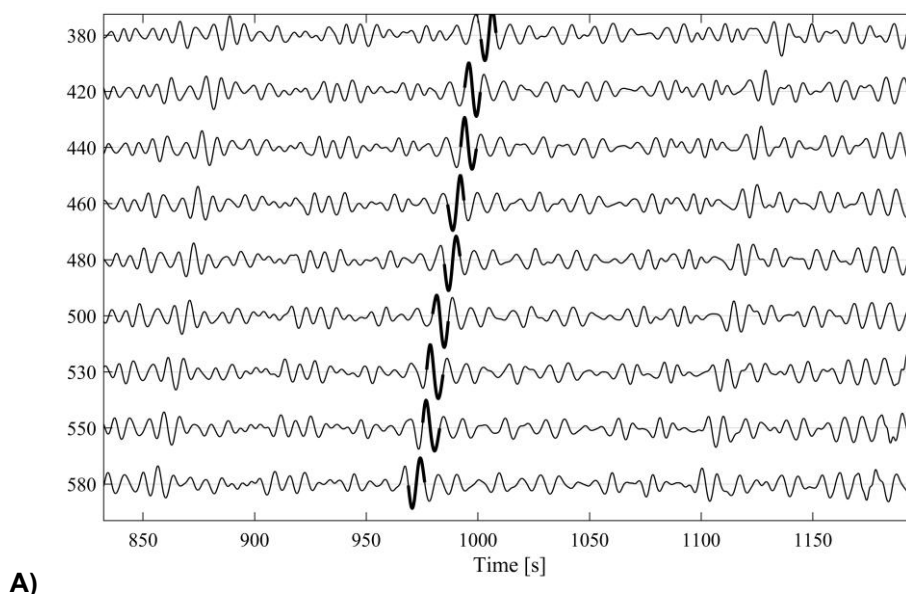
Суть реконструкции состоит в том, что эволюция волн рассчитывается в направлении берега (уменьшение x) от точки измерения и назад, в направлении моря на 1 км в рамках НУШ с переменными коэффициентами, как описано выше. Когда волны приближаются к береговой линии, их амплитуда возрастает, а в некоторых местах крутизна волн становится нереально большой. Обрушение волн не описывается в слабонелинейной модели, и волны могут расти бесконечно при приближении к берегу. Моделирование выполняется только на расстояния, когда параметр Урселла остается небольшим, – примерно до 150–250 м от берега. При значительных величинах параметра Урселла модель НУШ становится непригодной и должна быть заменена на модели Буссинесковского типа или уравнение Кортевега-де Вриза.

Данные численных расчетов сохранялись каждые 10 м; рассчитанные волновые поля длительностью 20 мин, охватывают пространственную область около 1.5 км. Поскольку локальная энергия волны изменяется из-за вариации глубины, значительная высота волны $H_s(x)$ рассчитывается из восстановленных временных последовательностей в каждой пространственной точке. Она оценивается в квазигауссовском приближении по стандартному среднеквадратическому отклонению смещения поверхности как $H_s = 4\eta_{rms}$. Восстановленная волна-убийца может исчезать на некотором расстоянии от x_0 , т.е. ее высота падает ниже уровня $2H_s(x)$, однако вскоре после этого может снова превысить этот порог, и поэтому естественно рассматривать вновь появляющиеся аномально большие

волны как одно непрерывное экстремальное событие. Как было найдено ранее для условий постоянной глубины [4], длительность таких событий, когда условие (1) может нарушаться в течение некоторых коротких периодов времени (менее пары периодов волн), может достигать 10 минут при глубоководном моделировании в рамках одномерных уравнений Эйлера.

На рис. 2а приведен пример реконструкции записи А. Волна-убийца возникает немного раньше точки регистрации ($x_0 = 500$ м) и распространяется примерно на 200 м в течение 33 секунд (приблизительно $4T_p$), при этом форма волны-убийцы в моделируемых волновых профилях изменяется в процессе эволюции. Подобные картины распространения получены и для случаев В и С. Время распространения экстремальных волн для этих записей составляет около $6T_p$ для случая В и $5T_p$ для случая С, расстояние, на которое распространяются волны-убийцы – 160 м и 260 м соответственно. Хотя временные последовательности в точке на рис. 2а демонстрируют, что волны-убийцы часто принадлежат к интенсивным группам волн, случаев, когда последующие волны в записи превышали бы значение $2H_s$ (т.е. формирование аномально высоких групп волн), не наблюдалось. Пространственно-временные (x, t) диаграммы смещения поверхности на рис. 2б позволяют выявить экстремальные события, а также определить времена их жизни. Точками на графике обозначены «одномоментные» волны-убийцы, определенные по критерию (1) и принадлежащие одному экстремальному событию (выделены квадратом). Все выделенные экстремальные события соответствуют волнам-убийцам, зарегистрированным в точке x_0 , и происходят на фоне интенсивных волновых пакетов, которые сохраняют свою энергию.

Полученные данные реконструкции – это порядка 150 временных рядов для каждого моделируемого события – позволяют оценить параметры волны-убийцы в статистическом смысле. Одним из таких параметров является ее геометрическая форма. Классификация форм на основе волновых асимметрий была введена в работе [4]. Мы различаем 4 геометрии волн-убийц: по вертикальной асимметрии (соотношение высоты гребня и глубины наибольшей ложбины) и по горизонтальной асимметрии (расположение наибольшей ложбины впереди или позади гребня). Например, Front Positive Wave (FPW) означает волну, у которой высота гребня больше, чем глубина впадины (потому Positive, а не Negative), а более глубокая ложбина расположена перед гребнем (потому Front, а не Rear).



A)

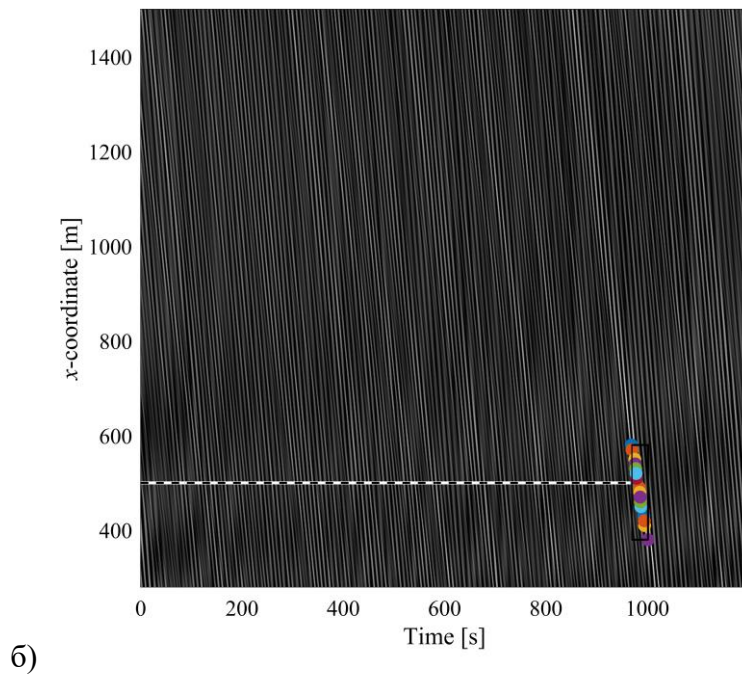


Рис. 2. Реконструкция эволюции волны-убийцы, серия А. а) Числа по вертикальной оси дают значения x в метрах; волна-убийца отмечена толстой линией; б) Пространственно-временные диаграммы смещения поверхности. Пунктирной линией обозначена точка измерения, символы соответствуют выделенным по критерию (1) волнам-убийцам.

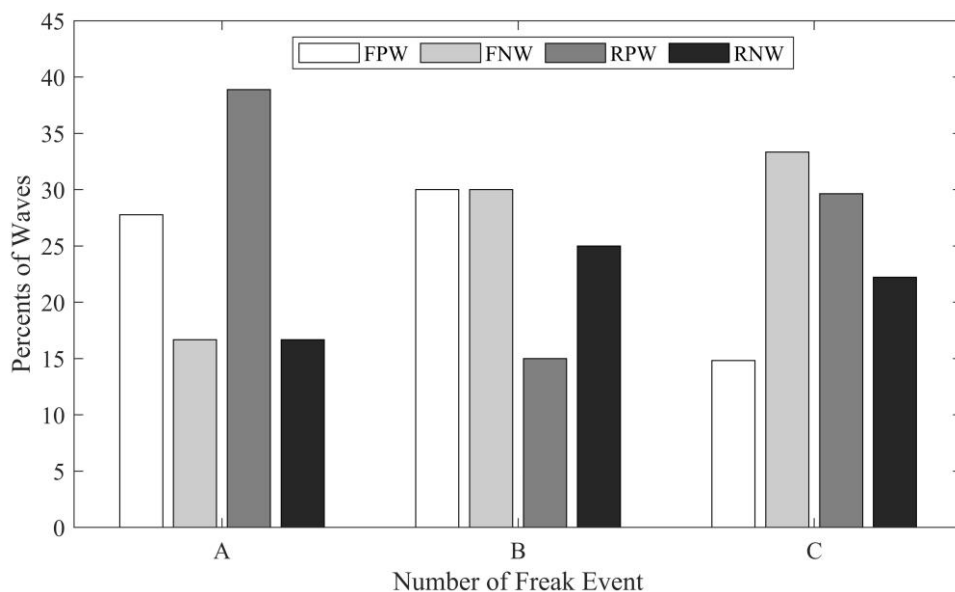


Рис. 3. Диаграммы распределения профилей волн-убийц по четырем характерным формам: 1) гребень выше ложбины, ложбина спереди глубже, чем сзади (Front Positive Wave); 2) гребень больше ложбины, ложбина сзади глубже, чем спереди (Rear Positive Wave); 3) ложбина больше гребня, ложбина спереди глубже, чем сзади (Front Negative Wave); 4) ложбина больше гребня, ложбина сзади глубже, чем спереди (Rear Negative Wave).

На рис. 3 представлена частота формирования этих четырех видов волн для каждой из трех записей. Количества волн оказывается недостаточно для получения статистически надежного результата, однако уже на основе этих данных можно говорить о том, что для случая А основная доля волн представлена положительными волнами с крутым задним

фронтом (Rear Positive Wave). Похожие результаты были получены нами ранее при реконструкции аномальных волн у берегов о. Тайвань, распространяющихся с глубокой воды на мелководье [2], а также в рамках стохастического моделирования уравнений Эйлера на глубокой воде [4]. Однако для серий В и С популяция аномальных волн с глубокими впадинами оказывается преобладающей. Для случая В выявлено больше волн с крутым передним фронтом (Front Wave), а для серии С доли волн с крутым передним или задним склоном (Front или Rear) примерно одинаковы.

Работа выполнена при поддержке гранта РФФИ 18-05-80019.

Библиографический список

1. Zeng H., Trulsen K. Evolution of skewness and kurtosis of weakly nonlinear unidirectional waves over a sloping bottom // Nat. Hazards Earth Syst. Sci. 2012. V. 12. P. 631–638.
2. Sergeeva A., Slunyaev A., Pelinovsky E., Talipova T., Doong D.-J. Numerical modeling of rogue waves in coastal waters // Nat. Haz. Earth Syst. Sci.. 2014. V. 14. P. 861–870.
3. Slunyaev A., Pelinovsky E., Guedes Soares C. Reconstruction of extreme events through numerical simulations // J. Offshore Mech. Arctic Eng. 2014. V. 136. P. 011302.
4. Sergeeva A., Slunyaev A. Rogue waves, rogue events and extreme wave kinematics in spatio-temporal fields of simulated sea states // Nat. Hazards Earth Syst. Sci. 2013. V. 13. P. 1759-1771.

A.V. Kokorina¹, E.G. Didenkulova^{1,2}

RECONSTRUCTION OF EXTREME SEA WAVES NEAR COASTAL WATERS OF SAKHALIN ISLAND

¹ *Institute of Applied Physics RAS, ² HSE University, Nizhny Novgorod*

Purpose: The research aims at reconstruction of rogue wave evolution on the sea surface by means of numerical simulations using the in-situ wave data.

Design/methodology/approach: The dynamics of extreme events was simulated using the variable-coefficient approximate evolution equations for weakly nonlinear waves, taking into account the shoaling effect.

Findings: The dynamics of rogue waves was reconstructed using the in-situ wave data measured in the coastal waters of Sakhalin Island. On the basis of numerical results, the characteristic shapes of anomalous waves, as well as the duration of their existence, were estimated.

Research limitations/implications: The application of employed methods is limited to weakly nonlinear and weakly dispersive case.

Originality/value: The research aims at solving one of the most important problems of oceanography – registration of the surface sea waves.

Keywords: gravity surface waves, nonlinear waves, in-situ measurements, numerical modeling, extreme sea wave.

**СРАВНИТЕЛЬНЫЙ АНАЛИЗ ТРЕХ ЛОГИСТИЧЕСКИХ МОДЕЛЕЙ
В ПРИЛОЖЕНИИ К МОДЕЛИРОВАНИЮ РАЗВИТИЯ ЭПИДЕМИИ COVID-19**¹*Нижегородский государственный технический университет им. Р. Е. Алексеева*²*Национальный исследовательский университет «Высшая школа экономики»*³*Институт прикладной физики РАН*

Эпидемия коронавируса нового типа COVID-19 оказала огромное негативное влияние на все области социальной, экономической и политической деятельности людей во всем мире. На сегодняшний день существует большое количество математических моделей для анализа и прогнозирования распространения коронавируса. В данной работе для описания динамики развития COVID-19 в 12 странах (Италия, Турция, Нидерланды, Южная Корея, Австрия, Швейцария, Испания, Бельгия, Германия, Чехия, Австралия, Норвегия) использованы три модели: простая (классическая) логистическая, обобщенная логистическая и модель Гомпертца. Проведен сравнительный анализ их эффективности. *Ключевые слова:* уравнение Гомпертца, логистическая модель, COVID-19.

Введение

В 2019 году мировое сообщество впервые столкнулось с инфекционным заболеванием COVID-19, вызываемым коронавирусом нового типа SARS-CoV-2 [1]. По данным Всемирной Организации Здравоохранения (ВОЗ) число, заразившихся в мире превысило сто миллионов человек, а число умерших – два миллиона (конец января 2021 года).

В связи с этим усилия большого количества специалистов по всему миру направлены на изучение данной коронавирусной инфекции и в том числе на построение адекватной математической модели, описывающей динамику ее поведения.

Для моделирования развития эпидемий и прогнозирования их последствий используется ряд математических моделей различной сложности. Наиболее известной моделью является логистическое уравнение Ферхюльста [2], представляющее собой нелинейное обыкновенное дифференциальное уравнение (ОДУ) первого порядка с постоянными коэффициентами. Качественно она объясняет зависимость числа переболевших от времени: экспоненциальный рост числа больных на начальном этапе развития эпидемии и выход на константу к ее завершению. В том случае, когда рост количества заболевших на начальном этапе степенной, в качестве модели можно использовать обобщенное логистическое уравнение [3, 4]. Очень часто для описания динамики развития инфекционных заболеваний применяют разностное логистическое уравнение или уравнение с запаздыванием [5, 6]. В более сложных моделях население разбивают на классы, например, «подверженные заболеванию», «заболевшие», «выздоровевшие» (так называемые SIR-модели), включающие в себя систему ОДУ высокого порядка. Динамика таких систем еще недостаточно изучена математически, и в ней возможны стохастические колебания [7]. Несмотря на это, данные модели довольно просты в реализации и показали свою эффективность для изучения динамики развития эпидемий.

Малопараметрические модели логистического типа достаточно хорошо описывают развитие эпидемии коронавируса, что подтверждается, например, в [8-13].

Наряду со «степенными» логистическими моделями в литературе достаточно давно была предложена «логарифмическая» модель Гомпертца, см., например, [14], которая практически не использовалась в эпидемиологических исследованиях. Эта модель обладает определенными преимуществами, поскольку содержит минимальное количество

параметров. В данной работе приведено описание модели Гомпертца, а также дан сравнительный анализ ее эффективности с простой и обобщенной логистической моделями при моделировании «первой волны» COVID-19.

Уравнение Гомпертца

Уравнение Гомпертца [14], записанное в стандартной форме, имеет вид:

$$\frac{dN}{dt} = rN \left(1 - \frac{\ln N}{\ln N_\infty} \right), \quad (1)$$

где r - скорость роста эпидемии и N_∞ - максимальное число переболевших людей.

Отметим, что число констант здесь такое же как в классическом логистическом уравнении – их две, и они имеют ясный смысл. Замечательным свойством этого уравнения является его интегрируемость и эквивалент логистической кривой есть

$$N = N_\infty e^{\frac{\ln \frac{N_0}{N_\infty} e^{-rt}}{N_\infty}}, \quad (2)$$

где N_0 – начальная численность заболевших коронавирусом.

Рассмотрим применимость этой модели для описания развития эпидемии COVID-19 на примере «первой волны» коронавирусной инфекции в 12 странах: Италия, Турция, Нидерланды, Южная Корея (далее ЮК), Австрия, Швейцария, Испания, Бельгия, Германия, Чехия, Австралия, Норвегия; исходные данные взяты с сайта Всемирной организации здравоохранения (<https://www.who.int/>). Данные даются с интервалом в один день, так что производную $dN/dt = K$ можно отождествить с количеством заболевших в один день. Тогда уравнение (1) можно рассматривать как алгебраическую связь между величинами N и K :

$$K = rN \left(1 - \frac{\ln N}{\ln N_\infty} \right), \quad (3)$$

Выбор стран обусловлен качеством исходных данных. На рис. 1 представлены данные медицинской статистики и алгебраической аппроксимации (3). Отметим, что кривая Гомпертца неплохо соответствует в среднем данным наблюдений, хотя суточное количество заболевших испытывает случайные вариации, которые могут быть связаны как с нарушением режима самоизоляции, так и с увеличением количества проводимых тестов.

Стандартный подход к определению коэффициентов модели состоит в минимизации среднеквадратического отклонения между модельным и реальным данными. Точность аппроксимации данных определяется с помощью коэффициента детерминации по формуле:

$$R^2 = 1 - \frac{\sum_{i=1}^n (K_i - \tilde{K}_i)^2}{\sum_{i=1}^n (K_i - \bar{K}_i)^2}, \quad (4)$$

где K_i – количество заболевших в i -й день, \tilde{K}_i – модельное значение, \bar{K}_i – среднее значение.

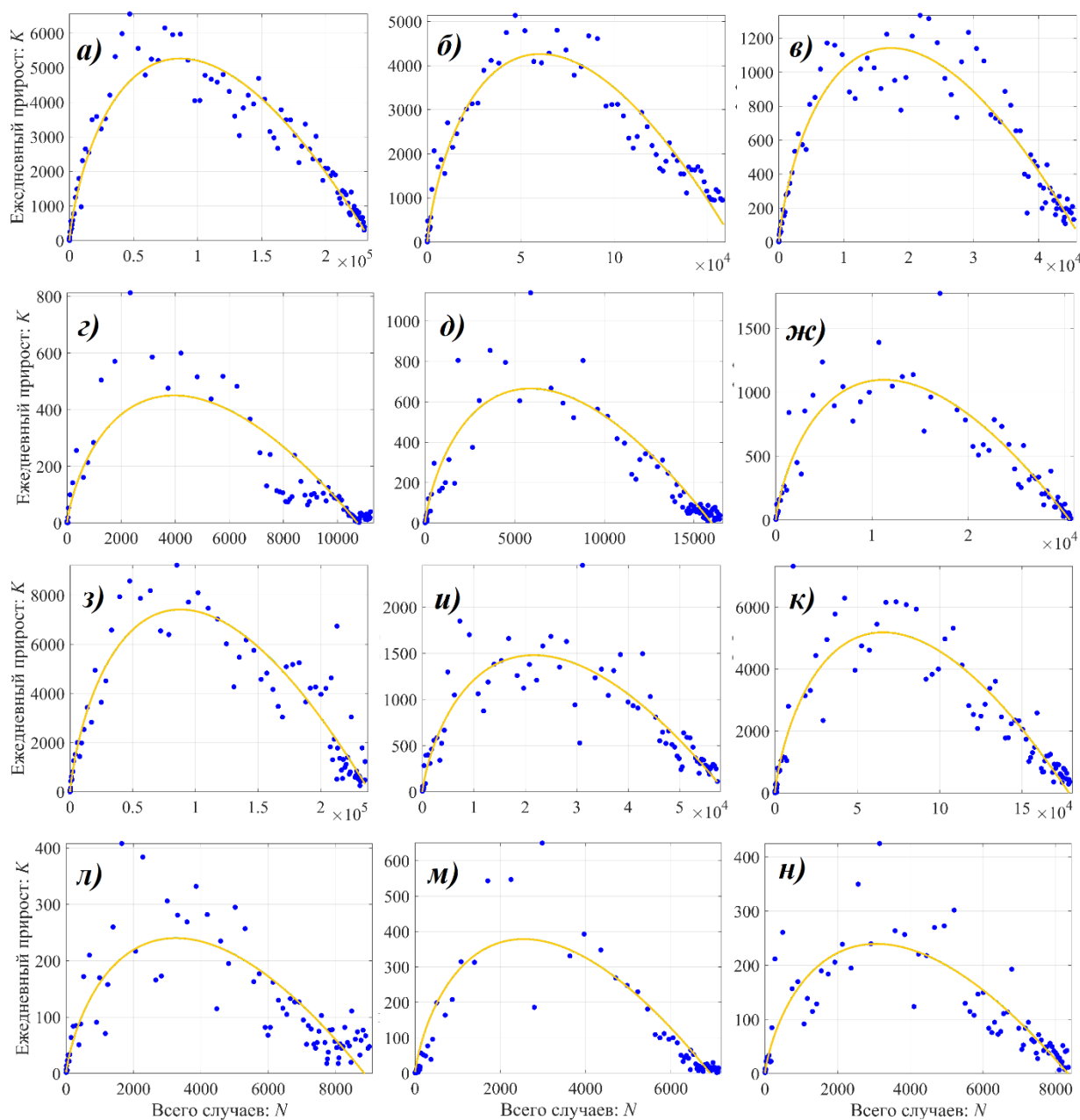


Рис. 1. Связь количества заболевших в сутки (K) и общего числа заболевших (N) в рамках модели Гомпертца (3) – сплошная линия. (а) Италия, (б) Турция, (в) Нидерланды, (з) Южная Корея, (д) Австрия, (е) Швейцария, (жс) Испания, (з) Бельгия, (и) Германия, (к) Чехия, (л) Австралия, (м) Норвегия

В таблице 1 приведены значения коэффициента детерминации R^2 для трех моделей: простой логистической модели (зеленый), её обобщения (красный) и модели Гомпертца (желтый). Результаты аппроксимаций в рамках логистической модели взяты из наших публикаций [12, 13]. Следует отметить, что для всех 12 стран коэффициент детерминации модели (3) значительно выше, чем для простой логистической модели, но несколько ниже, чем для обобщенной логистической модели (в которой есть четыре параметра для подгонки). В то же время число свободных параметров в классической логистической модели и модели Гомпертца одинаково (два параметра), так что можно сделать вывод, о преимуществе модели Гомпертца по сравнению с классической логистической моделью. В качестве примера на рис. 2 представлено сравнение аппроксимации тремя моделями на примере Австрии. Для наглядного представления результатов аппроксимации различными кривыми значения коэффициента R^2 приведено в виде гистограммы (рис. 3) для всех

стран. Среднее значение R^2 равно 0.66 для классической логистической модели, 0.88 для обобщенной логистической кривой и 0.8 для модели Гомпертца, а среднее квадратическое отклонение – 0.12, 0.07 и 0.09 соответственно. Очевидно, что для всех рассмотренных стран кривая Гомпертца достаточно точно описывает натурные данные, но, однако, она уступает обобщенной логистической модели.

Таблица 1

Коэффициенты модели и коэффициент детерминации для трёх аппроксимаций: зеленый – простая логистическая модель, красный – обобщенная логистическая и желтым представлена модель Гомпертца

Страна	N_∞	r	α	β	R^2	r	N_∞	R^2	N_∞	r	R^2
Италия	499300	0.994	0.8	5.1	0.96	0.749	253625	0.92	234800	0.12	0.73
Турция	146500	3.573	0.7	1.07	0.975	0.826	166042	0.89	161100	0.1	0.62
Нидерланды	41950	3.573	0.7	0.7	0.94	0.705	47099	0.904	45560	0.1	0.79
Южная Корея	16200	0.143	1.2	5.4	0.91	1.026	10829	0.72	10850	0.13	0.49
Австрия	18500	0.151	1.1	2.56	0.88	1.079	16155	0.82	15950	0.16	0.67
Швейцария	32100	1.093	0.8	1.5	0.86	1.005	30638	0.87	30230	0.14	0.76
Испания	256000	1.2	0.8	1.9	0.9	1.035	240385	0.86	234800	0.12	0.73
Бельгия	62000	2.2	0.7	1.3	0.9	0.743	59278	0.79	57350	0.1	0.66
Германия	211000	1.8	0.8	2.1	0.8	2.509	55828	0.63	176700	0.1	0.69
Чехия	14000	0.3	0.9	4.1	0.8	0.643	9045	0.64	8868	0.1	0.39
Австралия	8500	0.03	1.4	3.5	0.8	1.297	6974	0.81	6901	0.21	0.73
Норвегия	8900	0.6	0.9	1.6	0.8	0.701	8350	0.73	8182	0.114	0.6

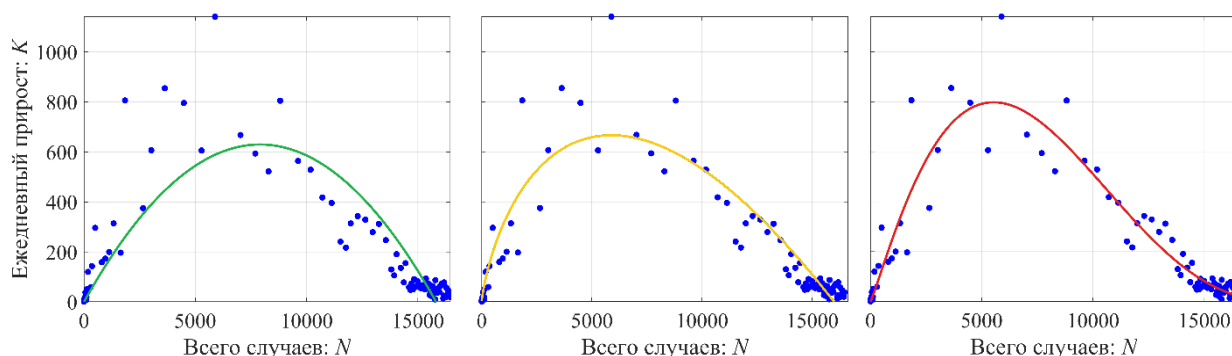


Рис. 2. Связь количества заболевших в сутки (K) и общего числа заболевших (N) для Австрии в рамках трех моделей: зеленый – простая логистическая модель, красный – обобщенная логистическая и желтым представлена модель Гомпертца

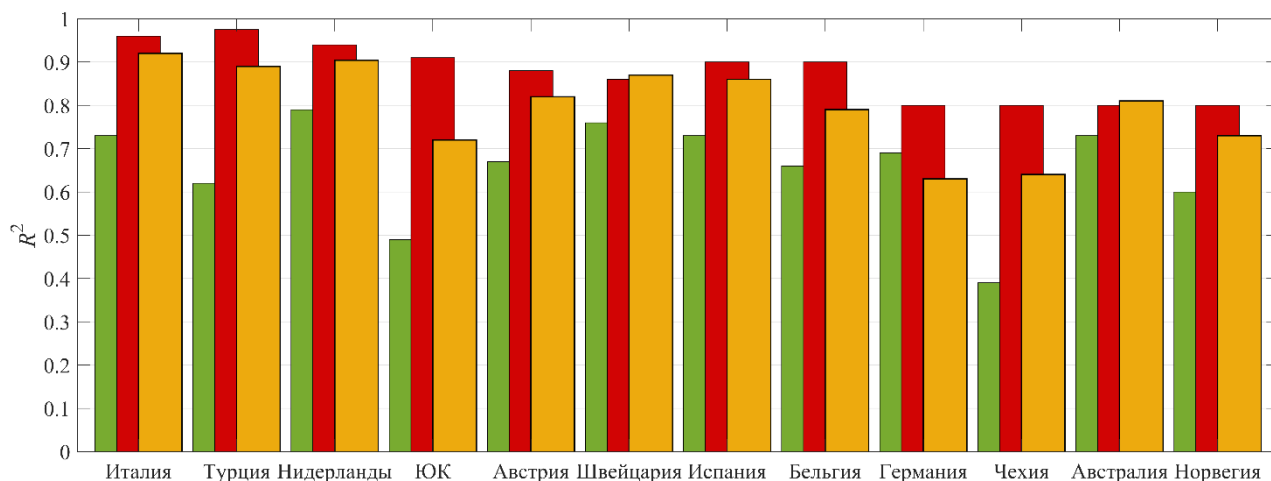


Рис. 3. Гистограмма распределения коэффициента детерминации R^2 по странам, цветом показана модель: зеленый – простая логистическая модель, красный – обобщенная логистическая и желтым представлена модель Гомпертца

Заключение

В данной работе мы проанализировали результаты математического моделирования динамики распространения «первой волны» новой коронавирусной инфекции COVID-19 на примере 12 стран в рамках трех различных моделей (классическая и обобщенная логистические модели и модель Гомпертца). Было показано, что модель Гомпертца описывает прирост заболевших от их общего количества заметно лучше, чем простая логистическая модель, но хуже, чем обобщенная.

Представленные результаты получены в рамках гранта РФФИ 20-51-80004 «Разработка мультимасштабных моделей для прогнозирования распространения КОВИД-19» для МК и АЕ, а также при поддержке Лаборатории динамических систем и приложений НИУ ВШЭ, грант Министерства науки и высшего образования РФ соглашение № 075-15-2019-1931 для ЕП.

Библиографический список

1. Ndaïrou F. Mathematical modeling of COVID-19 transmission dynamics with a case study of Wuhan / F. Ndaïrou, I. Area, J. Nieto, D. Torres // *Chaos, Solitons & Fractals*. – 2020. – V. 135. – Art.No. 109846. DOI: 10.1016/j.chaos.2020.109846.
2. Verhulst P. Notice sur la loi que la population suit dans son accroissement / P. Verhulst // *Corresp. Math. Phys.* – 1838. – T. 10. – С. 113-126. DOI: 10.12691/jbms-1-2-1
3. Blumberg A. A. Logistic growth rate functions / A. A. Blumberg // *Journal of Theoretical Biology*. – 1968. – V. 21. – № 1. – P. 42-44. DOI: 10.1016/0022-5193(68)90058-1
4. Brillhante M. F. Modeling risk of extreme events in generalized Verhulst models / M. F. Brillhante, M. I. Gomes, D. Pestana // *REVSTAT–Statistical Journal*. – 2019. – V. 17. – № 2. – P. 145-162.
5. Wright E. A non-linear difference-differential equation / E. Wright // *Journal für die reine und angewandte Mathematik*. – 1955. – T. 1955. – №. 194. – С. 66-87. DOI: 10.1515/crll.1955.194.66
6. Andrianov I. The Verhulst-Like Equations: integrable OΔE and ODE with Chaotic Behavior / I. Andrianov, G. Starushenko, S. Kvitka, L. Khajiyeva // *Symmetry*. – 2019. – V. 11. – № 12. – P. 1446. DOI: 10.3390/sym11121446.

7. Gray A. stochastic differential equation SIS epidemic model / A. Gray, D. Greenhalgh, L. Hu, X. Mao, J. Pan //SIAM Journal on Applied Mathematics. – 2011. – Т. 71. – №. 3. – С. 876-902. DOI: 10.1137/10081856X
8. Carletti T. COVID-19: The unreasonable effectiveness of simple models / T. Carletti, D. Fanelli, F. Piazza //Chaos, Solitons & Fractals: X. – 2020. – Т. 5. – С. 100034. DOI: 10.1016/j.csf.2020.100034
9. Consolini G. A stretched logistic equation for pandemic spreading / G. Consolini, M. Materassi //Chaos, Solitons & Fractals. – 2020. – Т. 140. – С. 110113. DOI: 10.1016/j.chaos.2020.110113
10. Wu K. Generalized logistic growth modeling of the COVID-19 outbreak: comparing the dynamics in the 29 provinces in China and in the rest of the world / K. Wu, D. Darcet, Q. Wang, D. Sornette //Nonlinear dynamics. – 2020. – Т. 101. – №. 3. – С. 1561-1581. DOI: 10.1101/2020.03.11.20034363
11. Куркин А. А. Логистические модели распространения эпидемий / А. А. Куркин, О. Е. Куркина, Е. Н. Пелиновский //Труды НГТУ им. ПЕ Алексева. – 2020. – №. 2 – С. 9-18. DOI: 10.46960/1816-210X_2020_2_9
12. Pelinovsky E. Logistic equation and COVID-19 / E. Pelinovsky, A. Kurkin, O. Kurkina, M. Kokoulina, A. Epifanova //Chaos, Solitons & Fractals. – 2020. – Т. 140. – С. 110241. DOI: 10.1016/j.chaos.2020.110241
13. Кокоулина М. В. Анализ динамики распространения коронавируса с помощью обобщенной логистической модели / М. В. Кокоулина, А. С. Епифанова, Е. Н. Пелиновский, О. Е. Куркина, А. А. Куркин//Труды НГТУ им. ПЕ Алексева. – 2020. – №. 3 – С. 28-49. DOI: 10.46960/1816-210X_2020_3_28
14. Gompertz B. On the nature of the function expressive of the law of human mortality and on the mode of determining the value of life contingencies / B. Gompertz // Philosophical Transactions of the Royal Society (London). - 1825. - V. 115. - P. 513-585. DOI:10.1098/rstl.1825.0026.

A.S. Epifanova¹, M.V. Kokoulina¹, O.E. Kurkina¹, A.A. Kurkin¹, E.N. Pelinovsky¹⁻³

COMPARATIVE ANALYSIS OF THREE LOGISTIC MODELS IN THE APPLICATION TO MODELING THE DEVELOPMENT OF THE COVID-19 EPIDEMIC

*Nizhny Novgorod state technical university n. a. R.E. Alekseev¹,
National Research University – Higher School of Economics²,
Institute of Applied Physics of the Russian Academy of Sciences³*

Purpose: to compare three logistic models in the application to modeling "first wave" development of COVID-19 in 12 countries.

Design/methodology/approach: we use logistic model, generalized logistic model and Gompertz model. The data on the incidence of COVID-19 are taken from the website of the World Health Organization.

Findings: It is shown that the Gompertz model describes the increase in cases from their total number better than the logistic model, but worse than the generalized model.

Originality/value: The obtained results may be used for the forecasting of the COVID-19 development.

Keywords: Gompertz equation, logistic model, COVID-19.

**ОНЛАЙН-РЕСУРС ДЛЯ ВИЗУАЛИЗАЦИИ И АНАЛИЗА ДИНАМИКИ
РАСПРОСТРАНЕНИЯ COVID-19 В РЕГИОНАХ РФ И ОТДЕЛЬНЫХ СТРАНАХ**

*Нижегородский государственный технический университет им. Р. Е. Алексеева¹,
Национальный исследовательский университет «Высшая школа экономики»²,
Институт прикладной физики РАН³, г. Нижний Новгород*

Пандемия коронавируса привела к худшему за последнее столетие экономическому кризису. Правительства по всему миру были вынуждены принять экстраординарные карантинные меры для замедления темпов распространения COVID-19. Усилия большого количества специалистов направлены на изучение коронавируса и, в том числе, на построение адекватной математической модели, описывающей динамику его распространения. В статье приведено описание онлайн-ресурса, предназначенного для визуализации и анализа тенденций развития эпидемии новой коронавирусной инфекции в 85 регионах РФ и 23 странах мира.

Ключевые слова: математическое моделирование, онлайн-ресурс, COVID-19

Введение

Новая коронавирусная инфекция впервые была обнаружена в китайском городе Ухань в декабре 2019 года. Уже к марту следующего года она достигла огромных масштабов, и Всемирная организация здравоохранения объявила о пандемии COVID-19. Пандемия коронавируса привела к худшему за последнее столетие экономическому кризису. Правительства по всему миру были вынуждены принять экстраординарные карантинные меры для замедления темпов распространения COVID-19: некоторые государства закрыли границы, во многих странах уже несколько раз вводился жесткий карантин и комендантский час, наложен запрет на проведение массовых мероприятий, большинство сотрудников переведено на удаленную работу, обязательным является использование средств индивидуальной защиты в общественных местах, людям, относящимся к категории риска, рекомендовано соблюдать режим самоизоляции. Число заразившихся в мире превысило 100 миллионов человек, а число умерших – два миллиона (по данным Всемирной организации здравоохранения на начало февраля 2021 года). В связи с этим усилия большого количества специалистов направлены на изучение данной коронавирусной инфекции и, в том числе, на построение адекватной математической модели, описывающей динамику ее распространения. Очевидно, что для разработки такой модели важную роль играет сбор исходных данных, их последующая визуализация и анализ.

Описание онлайн-ресурса

Коллективом Научно-исследовательской лаборатории Моделирования природных и техногенных катастроф НГТУ им. Р.Е. Алексеева создан сайт <https://lmpad.nntu.ru/>, на котором размещены сведения обо всех проектах сотрудников лаборатории, например IGWAtlas, IGWResearch, информация об экспериментах в гидродинамическом лотке и других (подробное их описание приведено в [1-3]). Один из проектов, открытый нами в самом начале пандемии, – ежедневно обновляемый онлайн-ресурс, содержащий информацию о динамике распространения COVID-19 во всех регионах России и 23 странах мира: США, Франции, Германии, Великобритании, Италии, Испании, Японии, Китае, Иране, Израиле, Южной Корее, Турции и ряде других государств. Изначально

страны были выбраны таким образом, чтобы они находились на разных этапах «первой волны» коронавируса.

Информация о заболеваемости новой коронавирусной инфекцией в России и мире обновляется ежедневно. Источник в РФ – интернет-ресурс для информирования населения по вопросам коронавируса (COVID-19) [4]. Сайт содержит официальную информацию по распространению коронавирусной инфекции в стране. Коммуникационный центр Правительства РФ предоставляет отчеты, которые включают, кроме прочего, сведения о количестве заболевших и умерших в мире, приросте заболевших в России и в регионах, количестве тестов на SARSCoV- 2, проведенных в РФ, статистике прибытия граждан в РФ, спросе на продовольственные товары, а также описание мер по противодействию COVID-19.

В качестве статистики по распространению коронавируса в мире используются данные с сайта Университета Джона Хопкинса [5], отчеты Всемирной организации здравоохранения [6].

Информационный ресурс, имеющий простой, интуитивно понятный пользовательский интерфейс, визуализирует данные в удобном для восприятия формате. Графики построены с помощью пакета прикладных программ MATLAB, который предоставляет пользователю большое количество библиотек для обработки и анализа данных, построения и визуализации математических моделей. Изображения (рис. 1) дают нам наглядное представление об эпидемиологической ситуации в каждом регионе РФ и в России в целом. Здесь и далее для всех графиков был выбран логарифмический масштаб вертикальной оси – в этом случае экспоненциальная зависимость будет характеризоваться прямой линией. В полулогарифмических координатах лучше видна скорость роста заболевших на начальном этапе. Штриховыми линиями обозначены экспоненциальные асимптоты, соответствующие удвоению числа случаев за определенное количество дней, звездочками – даты введения карантинных мер в России и других странах.

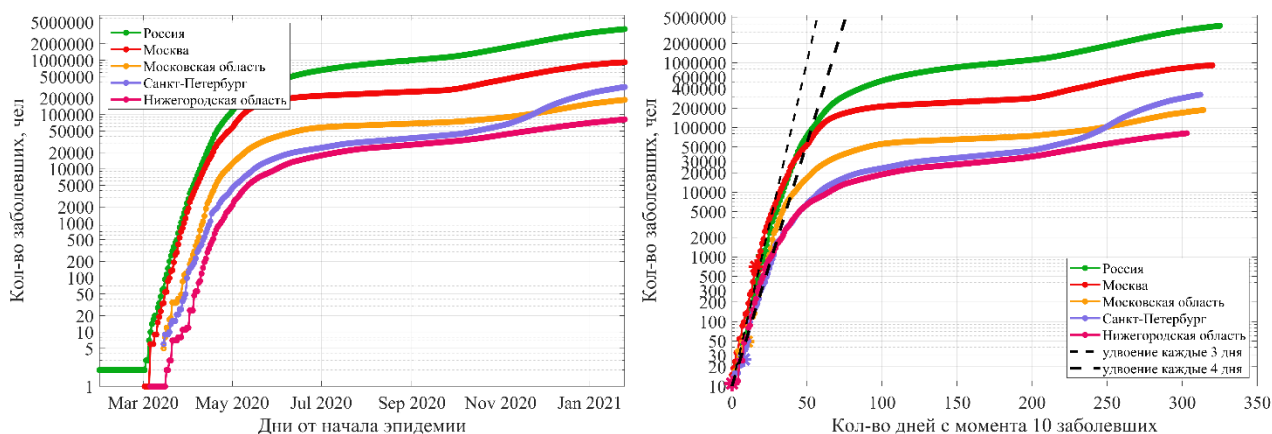


Рис. 1. Данные по России. Слева – график роста числа, заболевших от начала эпидемии (данные по датам), справа – график роста числа, заболевших с момента появления 10-ого заболевшего (данные по дням)

На первом месте по количеству подтвержденных случаев заболевания в России находится Москва, затем – Санкт-Петербург, Московская, Нижегородская и Свердловская области.

Динамика заболевания в некоторых странах мира проиллюстрирована на рис.2.

Очевидно, что США являются абсолютным мировым лидером по количеству подтвержденных случаев инфицирования: по информации Университета Джонса Хопкинса, в США за время пандемии зафиксировано более 27 млн. случаев заражения (на начало февраля 2021 года). На втором месте находится Индия (10,8 млн.), затем – Бразилия (9,5 млн.), Великобритания (3,98 млн.) и Россия (3,96 млн.).

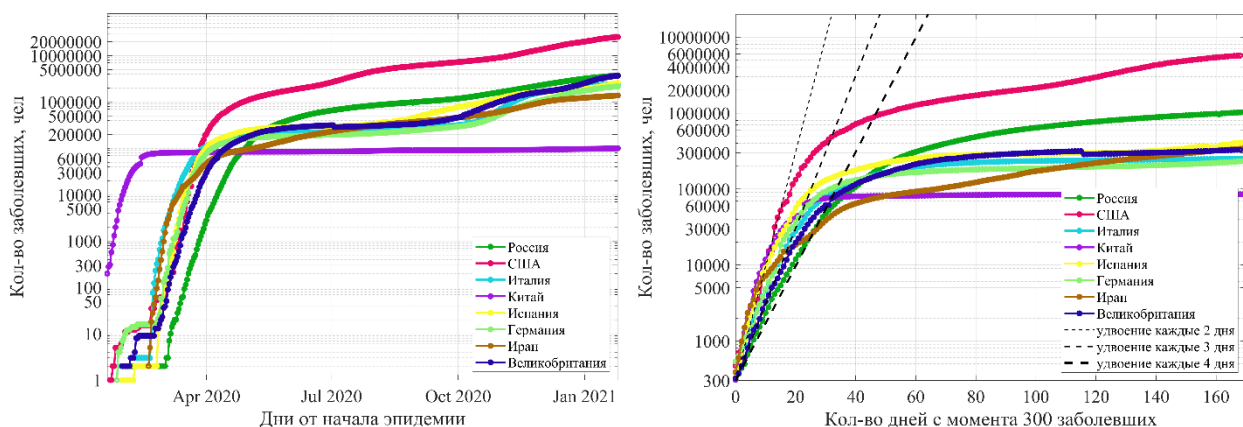


Рис. 2. Данные по миру. Слева – график роста числа, заболевших от начала эпидемии (данные по датам), справа – график роста числа, заболевших с момента появления 300-ого заболевшего (данные по дням).

Наиболее стабильной остается ситуация в Китае. Хотя впервые COVID-19 был зарегистрирован именно в КНР в конце 2019 года, и в начале марта 2021 года число подтвержденных случаев достигло 80 тыс., к настоящему времени в стране зарегистрировано чуть больше 100 тыс. случаев заражения.

На рис. 3 отображена тенденция роста эпидемии коронавирусной инфекции в России и мире. Точками обозначен прирост заболевших, выздоровевших и умерших в сутки, сплошными линиями – суммарные показатели за весь исследуемый период времени.

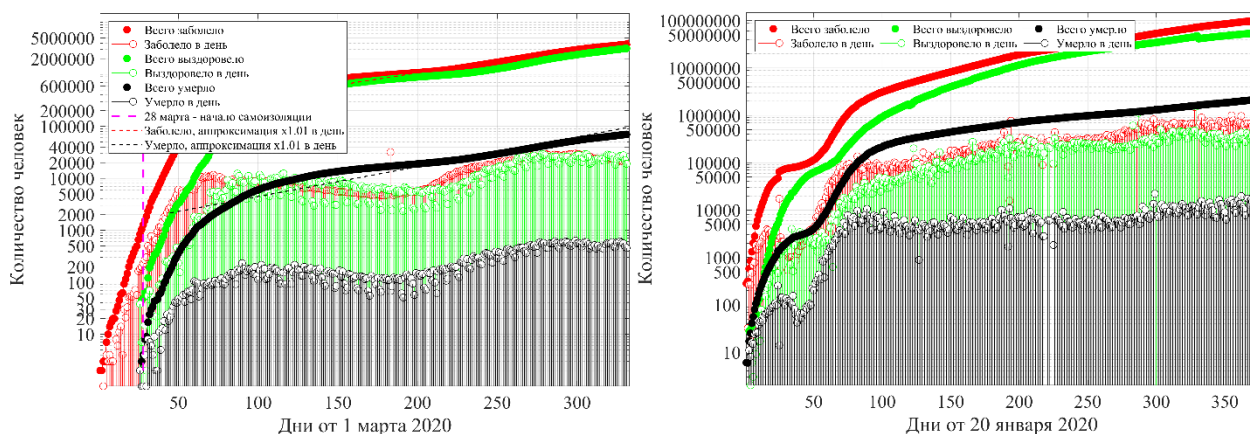


Рис. 3. Графики распространения эпидемии в России (слева) и в мире (справа)

Заключение

В статье приведено описание онлайн-ресурса, предназначенного для визуализации и анализа тенденций развития эпидемии новой коронавирусной инфекции. На сегодняшний день на сайте представлены графики, иллюстрирующие динамику инфицирования в 85 регионах РФ и 23 странах мира, а также прирост заболевших, выздоровевших и умерших в России и мире. Данные, агрегируемые сайтом, используются для построения логистических моделей развития эпидемии. Ознакомиться с ними более подробно можно в [7-9].

Представленные результаты получены при поддержке гранта Президента РФ по государственной поддержке ведущих научных школ РФ НШ-2485.2020.5, а также при поддержке РФФИ и [CNPq, andNSFC] в рамках проекта № 20-51-80004.

Библиографический список

1. Епифанова А.С. База данных лабораторных экспериментов по генерации и распространению внутренних и поверхностных волн. Первые результаты / А.С. Епифанова, О.Е. Куркина, Н.М. Лиходеев, А.Р. Гиниятуллин, Н.А. Санников, Л.В. Талалушкина // Сборник материалов 30-й Всероссийской научно-практической конференции по графическим информационным технологиям и системам. – 2020. – С. 302-306.
2. Куркин А.А. Лабораторное и численное моделирование стратифицированных потоков с геофизическими приложениями / А.А. Куркин, О.Е. Куркина, Л.В. Талалушкина и др. // Моря России: фундаментальные и прикладные исследования. – 2019. – С. 88-90.
3. Тюгин Д.Ю. Проблемно-ориентированный программный комплекс для моделирования динамики внутренних волн в стратифицированном океане / Д.Ю. Тюгин, А.А. Куркин, О.Е. Куркина // Труды НГТУ им. Р.Е. Алексева. – 2018. – № 2. – С. 52-59.
4. Куркин А. А. Логистические модели распространения эпидемий / А. А. Куркин, О. Е. Куркина, Е. Н. Пелиновский // Труды НГТУ им. Р.Е. Алексева. – 2020. – №. 2 – С. 9-18. DOI: 10.46960/1816-210X_2020_2_9.
5. Коммуникационный центр Правительства Российской Федерации [Электронный ресурс]. – Режим доступа: <https://стопкоронавирус.рф/>, свободный (дата обращения: 01.02.2021).
6. The Johns Hopkins Coronavirus Resource Center [Электронный ресурс]. – Режим доступа: <https://coronavirus.jhu.edu/map.html>, свободный (дата обращения: 01.02.2021).
7. WHO Coronavirus Disease (COVID-19) Dashboard [Электронный ресурс]. – Режим доступа: <https://covid19.who.int/table>, свободный (дата обращения: 01.02.2021).
8. Pelinovsky E. Logistic equation and COVID-19 / E. Pelinovsky, A. Kurkin, O. Kurkina, M. Kokoulina, A. Epifanova // Chaos, Solitons & Fractals. – 2020. – Т. 140. – С. 110241. DOI: 10.1016/j.chaos.2020.110241.
9. Кокоулина М. В. Анализ динамики распространения коронавируса с помощью обобщенной логистической модели / М. В. Кокоулина, А. С. Епифанова, Е. Н. Пелиновский, О. Е. Куркина, А. А. Куркин // Труды НГТУ им. Р.Е. Алексева. – 2020. – №. 3 – С. 28-49. DOI: 10.46960/1816-210X_2020_3_28.

A.S. Epifanova¹, M.V. Kokoulina¹, O.E. Kurkina¹, A.A. Kurkin¹, E.N. Leshekhva,
E.N. Pelinovsky¹⁻³

ONLINE RESOURCE FOR VISUALIZATION AND ANALYSIS OF THE DYNAMICS OF THE COVID-19 DISTRIBUTION IN THE REGIONS OF THE RUSSIAN FEDERATION AND SOME COUNTRIES

Nizhny Novgorod state technical university n. a. R.E. Alekseev¹, National Research University – Higher School of Economics², Institute of Applied Physics of the Russian Academy of Sciences³

Purpose: creation of a daily updated online resource containing information on the dynamics of the COVID-19 development.

Design/methodology/approach: data from the websites of the RF Government Communication Center, Johns Hopkins University, and the World Health Organization are used as statistics on the spread of coronavirus. The graphs were built using MATLAB application software package.

Findings: the online resource is designed to visualize and analyze trends in the epidemic of new coronavirus infection in 85 regions of the Russian Federation and 23 countries around the world.

Originality/value: the online resource helps in the study of COVID-19 dynamics.

Keywords: mathematical modeling, online resource, COVID-19.

**РАСПРОСТРАНЕНИЕ ЛОКАЛИЗОВАННОГО ПАКЕТА ВНУТРЕННИХ ВОЛН
НАД БАТИМИТРИЧЕСКОЙ НЕОДНОРОДНОСТЬЮ***Нижегородский Государственный Технический Университет им. Р. Е. Алексеева*

Данная работа посвящена изучению особенностей трансформации волнового локализованного пакета над донным уступом в жидкости с симметричной трехслойной стратификацией плотности в рамках уравнения Гарднера. В настоящей работе исследование проведено для различных значений высоты и ширины донного уступа, при фиксированном начальном возмущении в виде бризероподобного пакета внутренних волн. Продемонстрированы начальные различные сценарии трансформации волнового пакета. Проведены количественные амплитудные оценки волнового поля после уступа. Показано, что при прохождении в зону меньшей глубины амплитуда волн всегда уменьшается вследствие сложной зависимости нелинейности от параметров среды в модели. Были рассчитаны статистические моменты до четвертого порядка. Их зависимость от пространственной координаты дает представление о волновом режиме в различных областях.

Ключевые слова: внутренние волны, слабонелинейная теория, бризер, трехслойная стратификация, уравнение Гарднера, неоднородная среда, донный уступ.

Введение

Внутренние волны довольно распространенное явление, они часто наблюдаются в стратифицированных водоемах: океане [1] и озерах [2]. Внутренние волны распространяются в зоне шельфа – континентального склона и играют важную роль в ресуспендировании наносов и каскадах энергии [3], а также оказывают важное влияние на окружающую среду [4]. Считается, что на экологические системы и качество воды в океанах и озерах влияет долгосрочный перенос массы, связанный с внутренними волнами [5]. Чтобы лучше понять эти процессы, необходимо выяснить, какие типы внутренних волн могут генерироваться как они распространяются при различных типах неровностей дна. Настоящая работа является продолжением исследований, начатых в работе [6]. Объектом исследования являются локализованные пакеты внутренних волн (бризеры) или «осциллирующие пульсирующие солитоны» («oscillatory pulse solitons») [7]. В настоящее время бризеры изучены довольно слабо, поскольку их трудно определить в записях волновых полей и идентифицировать в поле внутренних волн. Здесь представлено исследование динамики бризеров в неоднородной среде, дополняющее круг задач эволюции и трансформации внутренних волн в рамках моделей слабонелинейной теории [8,9].

Волновые режимы в рамках уравнения Гарднера

В данной работе будем рассматривать процесс трансформации начальных внутренних волновых возмущений над донным уступом в рамках нелинейного эволюционного уравнения Гарднера (или модифицированного уравнения Кортевега – де Вриза (мКдВ), в случае, когда коэффициент квадратичной нелинейности равен нулю) [10]. Уравнение мКдВ при $\alpha_1 > 0$ обладает точными стационарными (солитонами) и квазистационарными (бризеры) решениями (см. Рис. 1). Хотя бризеры и представляют собой локализованные возмущения, распространяющимися без потерь, они имеют внутренние осцилляторные степени свободы и не являются волнами неизменной формы. При отрицательном коэффициенте кубической нелинейности ($\alpha_1 < 0$) решений в виде бризеров не существует, однако существует семейство солитонов, полярность которых определяется знаком коэффициента квадратичной нелинейности. Все возможные формы

решения уравнения Гарднера при различных знаках коэффициентов квадратичной и кубической нелинейности представлены на рисунке 1.

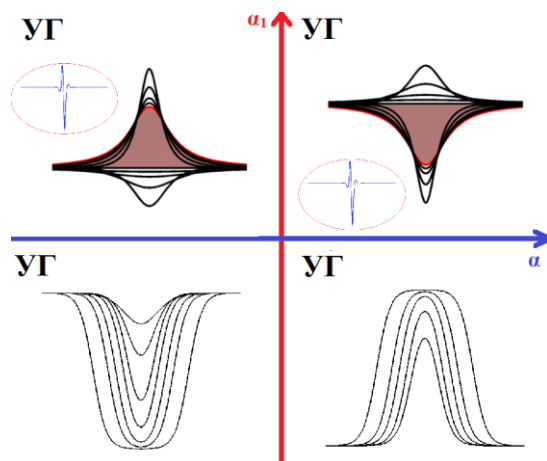


Рис. 1. Возможные формы решений уравнения Гарднера при различных знаках коэффициентов квадратичной и кубической нелинейности [6]; Горизонтальная ось соответствует уравнению КдВ ($\alpha_1=0$); Вертикальная ось – уравнению мКдВ ($\alpha = 0$); УГ – уравнение Гарднера

На основании приведенной плоскости параметров (см. Рис. 1) можно лишь качественно оценить вид возможных волновых возмущений при трансформации бризера над уступом различной ширины и высоты, поскольку изменения значений коэффициентов всегда можно рассчитать для при известных параметрах среды, например, для трехслойной стратификации плотности формулы даны в работе [11]. Для более точных количественных оценок их характеристик в горизонтально-неоднородной жидкости необходимо проводить численные расчеты, учитывая свойства теоретических решений при анализе.

Постановка задачи

В рамках данной работы ставится задача об изучении особенностей трансформации бризера над донным уступом различной высоты и ширины в симметричной трехслойной стратификации жидкости. На рисунке 2 изображена схема расчетной области и начального возмущения поля плотности в виде бризера уравнения мКдВ с амплитудой $a = 14$ м, начальные параметры бризера аналогичны [6]. Длина исследуемой области принималась равной $L = 60000$ м, невозмущенная глубина до уступа $-H = 100$ м с высотой уступа H_s и характерной шириной W_s (см. Табл. 1). Форма сглаженного уступа определялась следующим соотношением:

$$D(x) = \frac{-(H - H_s)}{2} - \frac{H_s}{2} \operatorname{th} \frac{x - x_0}{d},$$

где $x_0 = 25000$ м – положение центра зоны неоднородной глубины по оси x , $W_s = 2d$. Сглаженная стратификация плотности с двумя пикноклинами задается следующим соотношением:

$$\rho(z) = \rho_0 - \Delta\rho_1 \operatorname{th} \frac{z - z_{\text{пк1}}}{d_{\text{пк1}}} - \Delta\rho_2 \operatorname{th} \frac{z - z_{\text{пк2}}}{d_{\text{пк2}}},$$

где $\rho = 1020$ кг/м³ – среднее значение плотности, $\Delta\rho = \Delta\rho_1 = \Delta\rho_2 = 5$ кг/м³ – скачки плотности на верхнем и нижнем пикноклинах соответственно, $z_{\text{пк1}} = -30$ м, $z_{\text{пк2}} = -70$ м – глубины залегания пикноклинов, $d_{\text{пк1}} = d_{\text{пк2}} = 4$ м – характерная полуширина пикноклинов.

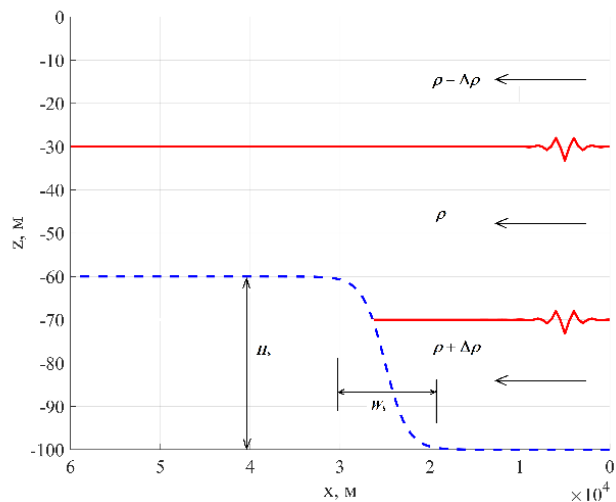


Рис. 2. Схема задачи

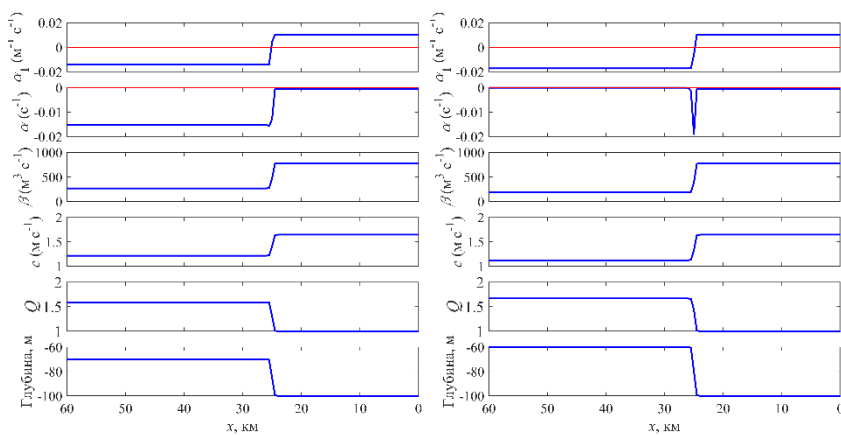
Таблица 1

Параметры неоднородности дна в численных экспериментах

$\frac{H_s}{W_s}$	30 м	40 м	50 м
500 м	Э1	Э2	Э3
5000 м	Э4	Э5	Э6
10000 м	Э7	Э8	Э9

Основные результаты

Моделирование производилось для уступа с параметрами, приведенными в таблице 1. В зависимости от сочетания этих параметров, изменяются значения коэффициентов уравнения Гарднера. На рисунке 3 приведены коэффициенты в серии численных экспериментов Э1, Э4, Э7 (см. Табл. 1).



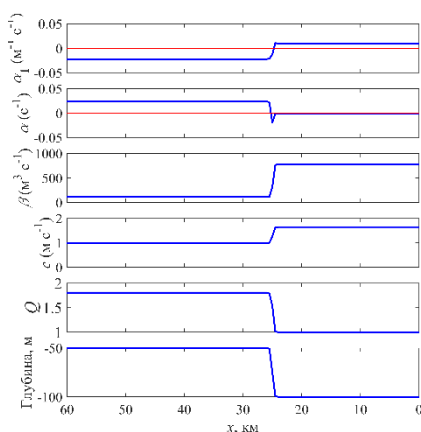


Рис. 3. Коэффициенты уравнения Гарднера вдоль трассы распространения в серии численных экспериментов Э1, Э4, Э7

Коэффициент кубической нелинейности остается отрицательным во всей числовой области в экспериментах Э1- Э9, а коэффициент квадратичной нелинейности меняет свой знак в зависимости от высоты донного уступа (см. Табл. 2). Для уступа высотой 40 м. (эксперименты Э4 – Э6) коэффициент квадратичной нелинейности возвращается к нулевому значению, так как при такой высоте уступа в зоне горизонтальной неоднородности наблюдается симметричная двухслойная стратификация плотности.

Таблица 2

Знак коэффициента квадратичной нелинейности после уступа для различных численных экспериментов

Маркер эксперимента	Э1	Э2	Э3	Э4	Э5	Э6	Э7	Э8	Э9
Знак α	$\alpha < 0$	$\alpha < 0$	$\alpha < 0$	$\alpha = 0$	$\alpha = 0$	$\alpha = 0$	$\alpha > 0$	$\alpha > 0$	$\alpha > 0$

Как известно для анализа волновых полей широко используются статистические моменты [12], которые вычисляются следующим образом:

$$\mu_n(t) = \int_{-\infty}^{+\infty} \eta^n(x, t) dx,$$

где, $\eta(x, t)$ - решение уравнения Гарднера [6], n – порядок статистического момента.

Первый и второй моменты соответственно равны математическому ожиданию и среднеквадратичному отклонению, и над ровным дном постоянны во времени, а третий и четвертый моменты соответствуют асимметрии и эксцессу, то есть характеристикам случайных процессов. Подробный вывод формул и теоретическое описание характеристик и свойств статистических моментов для частных случаев можно найти в работе [13]. На рисунке 4 приведены первые четыре статических момента для трансформации бризера в эксперименте Э7.

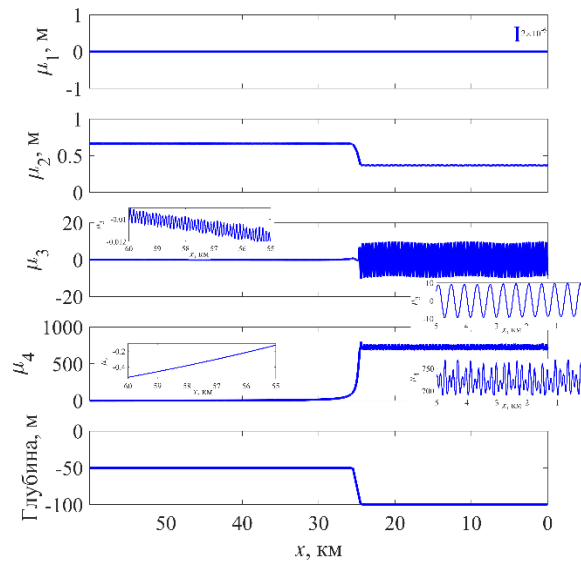


Рис. 4. Статистические моменты. Эксперимент Э7.

Можно отметить, что в силу свойств бризерного решения уравнения мКдВ (нулевая «масса») первый статистический момент равен нулю (см. Рис. 4). Третий и четвертый моменты до уступа характеризуются периодическими изменениями в пространстве, что было отмечено в работе [13]. Размах изменений довольно велик, это является отличительным свойством бризеров. При прохождении бризером зоны неоднородности они резко уменьшаются по величине (см. Рис. 4), что свидетельствует о значительной трансформации волнового пакета над уступом и потере его идентичности с преобразованием в другие волновые режимы.

На рисунке 5 приведены пространственно-временные диаграммы процесса трансформации бризера в зоне неоднородности в зависимости от высоты уступа в серии экспериментов Э1, Э4, Э7 (см. Табл. 1).

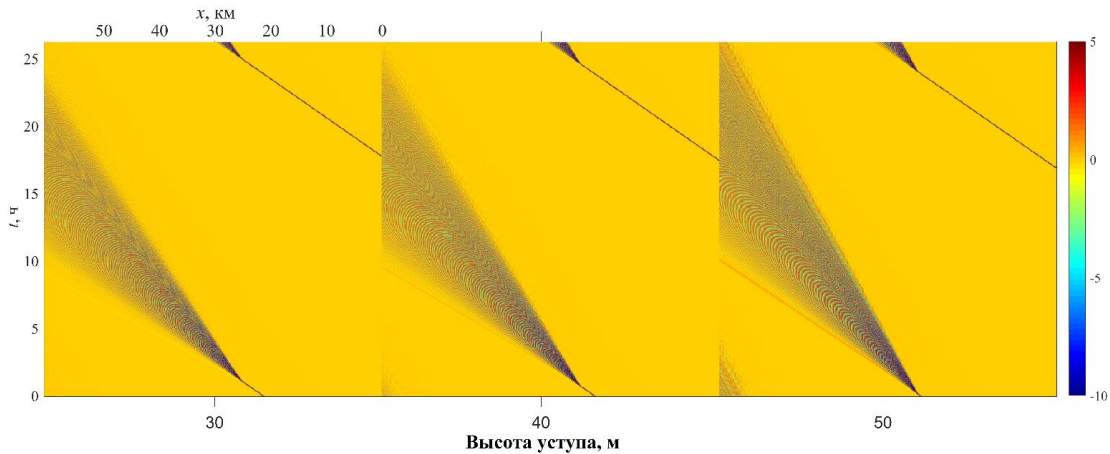


Рис. 5. Пространственно-временные диаграммы процесса трансформации бризера над донным уступом. Серия экспериментов Э1, Э4, Э7 (слева направо, см. Табл. 1)

В приведенных расчетах во всех случаях наблюдается значительное уменьшение амплитуды нелинейного волнового пакета с образованием диспергирующего цуга волн, амплитуда которого в приведенных случаях составляет приблизительно 28% от амплитуды исходного волнового пакета. При этом необходимо отметить, что в эксперименте Э1 наблюдается генерация уединенной внутренней волны отрицательной полярности, амплитуда которой составляет 2% от начального возмущения (см. Рис. 6, сверху). В эксперименте Э7 наблюдается генерация волны положительной полярности с амплитудой 5% от исходной амплитуды бризера (см. Рис. 6, снизу), что согласуется с качественными теоретическими оценками, приведенными выше (см. Рис. 1).

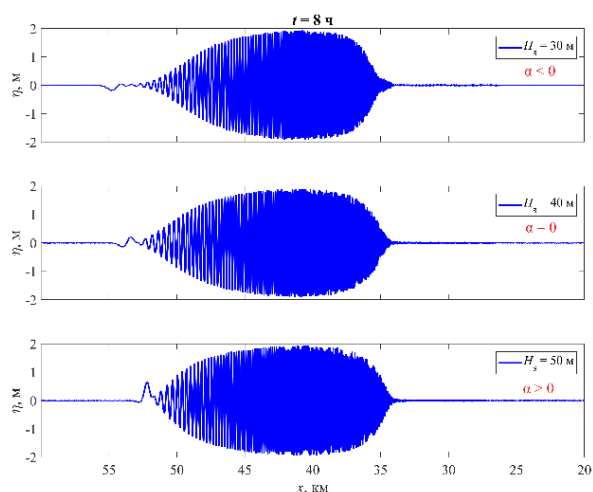


Рис. 6. Смещение верхнего пикноклина в момент времени $t = 8$ ч. Серия экспериментов Э1, Э4, Э7 (сверху вниз, см. Табл. 1)

Заключение

Проведен цикл численных расчетов по исследованию процесса распространения нелинейного волнового пакета над горизонтальной неоднородностью с различными начальными параметрами в рамках уравнения Гарднера. Были построены и проанализированы пространственно-временные диаграммы процесса трансформации бризера и рассчитаны статистические моменты до четвертого порядка. Проведены количественные амплитудные оценки волнового поля прошедших волн. Показано, что при прохождении в зону меньшей глубины амплитуда волн всегда уменьшается вследствие сложной зависимости нелинейности от параметров среды в модели. Зависимость статистических моментов от пространственной координаты дает представление о волновом режиме до и после донного уступа. Полученные численные результаты согласуются со слабонелинейной теорией внутренних волн.

Представленные результаты получены при поддержке гранта Президента РФ по государственной поддержке ведущих научных школ РФ НШ-2485.2020.5 и молодых российских ученых – кандидатов наук МК-218.2020.5

Библиографический список

1. Lamb K. G. Internal wave breaking and dissipation mechanisms on the continental slope/shelf // Annual Review of Fluid Mechanics. 2014. Т. 46. С. 231-254.
2. Wüest A., Lorke A. Small-scale hydrodynamics in lakes // Annual Review of fluid mechanics. 2003. Т. 35. №. 1. С. 373-412.
3. Boegman L., Stastna M. Sediment resuspension and transport by internal solitary waves // Annual review of fluid mechanics. 2019. Т. 51. С. 129-154.
4. Woodson C. B. The fate and impact of internal waves in nearshore ecosystems // Annual review of marine science. 2018. Т. 10. С. 421-441.
5. Alford MH. Redistribution of energy available for ocean mixing by long-range propagation of internal waves // Nature. 2003. Vol.423 P.159–162
6. Талалушкина Л.В., Лобовиков П.В., Кокоулина М.В., Гиниятуллин А.Р. Трансформация внутреннего бризера над уступом в рамках уравнения Гарднера // Сборник материалов XXVI Международной научно-технической конференции «Информационные системы и технологии ИСТ-2020». 2020. С. 1198-1204
7. Куркина О.Е., Рувинская Е.А., Соомере Т., Талипова Т.Г. Нелинейные пакеты внутренних волн: динамика бризеров // Труды всероссийской конференции «Прикладные технологии гидроакустики и гидрофизики» Федеральное

- государственное бюджетное учреждение науки Институт океанологии им. П.П. Ширшова. Российской академии наук. 2016. №. 13.С. 170-172.
8. Wadati M. The Modified Korteweg-deVries Equation. // Jour. Phys. Society Jpn. 1973. V.34. P.1289
 9. Grimshaw R., Pelinovsky E., Talipova T. The modified Korteweg–de Vries equation in the theory of the large–amplitude internal waves. //Nonlinear Processes in Geophysics. 1997. V. 4. P. 237–250.
 10. Полухина О.Е., Пелиновский Е.Н., Слюняев А.В. Обобщенное уравнение Гарднера для внутренних волн в стратифицированной жидкости // Препринт ИПФ РАН. 2002 № 606. 28 с.
 11. Kurkina O., Rouvinskaya E., Talipova T. & Soomere T. Propagation regimes and populations of internal waves in the Mediterranean Sea basin. // Estuarine, Coastal and Shelf Science. 2017. V. 185, P. 44–54.
 12. Monin A.S. Yaglom A.M. Statistical Fluid Mechanics // MIT Press: Cambridge, MA, USA, 1971. P. 769
 13. Didenkulova E., Pelinovsky E. Breather’s Properties within the Framework of the Modified Korteweg–de Vries Equation //Symmetry. 2020. Vol. 12. №. 4. P. 638

L. Talalushkina, P. Lobovikov, O. Kurkina, A. Giniyatullin

PROPAGATION OF A LOCALIZED PACKET OF INTERNAL WAVES OVER A BATHYMETRIC INHOMOGENEITY

Nizhny Novgorod state technical university n. a. R. E. Alekseev

This work is devoted to the study of the features of the transformation of a localized wave packet over a bottom step in a liquid with symmetric three-layer density stratification within the framework of the Gardner equation. In this work, the study was carried out for various values of the height and width of the bottom step, with a fixed initial disturbance in the form of a breather-like packet of internal waves. Various initial scenarios of wave packet transformation are demonstrated. Quantitative amplitude estimates of the wave field after the step are carried out. It is shown that when passing into a zone of shallow depth, the wave amplitude always decreases due to the complex dependence of nonlinearity on the parameters of the medium in the model. Statistical moments up to the fourth order were calculated. Their dependence on the spatial coordinate gives an idea of the wave regime in different regions.

Keywords: internal waves, weakly nonlinear theory. breather, three-layer fluid, Gardner equation, bottom step

ЭПИЗОДЫ НАБЛЮДЕНИЙ БРИЗЕРОПОДОБНЫХ ВНУТРЕННИХ ВОЛН В МИРОВОМ ОКЕАНЕ

Нижегородский государственный технический университет им. Р.Е. Алексеева

Целью настоящей работы является сбор и систематизация информации о наблюдениях бризероподобных внутренних волновых пакетов в Мировом океане и определение акваторий, где теоретически возможно проявление изучаемых явлений. В работе проанализированы пять записей внутренних волн, имеющихся в базе данных IGWAtlas, на предмет возможного проявления в них бризероподобных волновых пакетов.

Ключевые слова: натурные данные, внутренние волны, бризеры, база данных

Введение

Натурные данные (как прямых, так и дистанционных наблюдений) подтверждают возникновение и существование внутри океана интенсивных короткопериодных внутренних волн, которые переносят значительную энергию на большие расстояния и существенным образом влияют на окружающую их среду (от перераспределения по вертикали питательных веществ и размывов дна до искажения передаваемых на дальние расстояния акустических сигналов).

Бризеры, или нелинейные локализованные волновые пакеты, представляют собой важный класс неизлучающих квазистационарных объектов в динамике нелинейных волновых систем. Исследование бризеров в теории внутренних волн начато относительно недавно, теоретически и численно [Grimshaw et al., 1999; Lamb et al., 2007; Rouvinskaya et al., 2015]. Здесь нужно отметить сложность их идентификации в записях внутренних волн, где они трудно отличимы от линейных пакетов. Несмотря на это, начинают появляться новые наблюдения бризеров как в экспериментах [Zhang et al., 2018], так и в природных условиях [Lee et al., 2006; Shroyer et al., 2010; Кокоулина и др., 2020]. Например, в работе [Кокоулина и др., 2020] отмечается наличие похожих на бризеры волновых пакетов на записях полей температуры в Японском море.

В отличие от солитонов, которые часто и повсеместно наблюдаются в зонах генерации внутренних волн и на шельфах, локализованным нелинейным волновым пакетам (бризерам) уделялось очень мало внимания. Поскольку их трудно идентифицировать в поле внутренних волн, в литературе можно найти лишь немногочисленные эпизоды наблюдений, содержащих бризероподобные пакеты.

Наблюдения бризероподобных волновых пакетов

Коллективом научно-исследовательской лаборатории Моделирования природных и техногенных катастроф НГТУ им. Р.Е. Алексеева разработано веб-приложение IGWAtlas, которое позволяет взаимодействовать с базой данных (БД) наблюдений внутренних волн в Мировом океане (более подробно IGWAtlas и БД описаны в [Epifanovaetal., 2019]). В рамках данной статьи проанализированы пять описанных в литературных источниках эпизодов наблюдений внутренних волн [Lee et al., 2006; Shroyer et al., 2010; Кокоулина и др., 2020, Masunaga et al., 2016, Ostrowski&Bazika-Sangolay, 2015]. (рис. 1). Предполагается, что все они являются бризероподобными волновыми пакетами. Рассмотрим подробнее каждый из эпизодов.

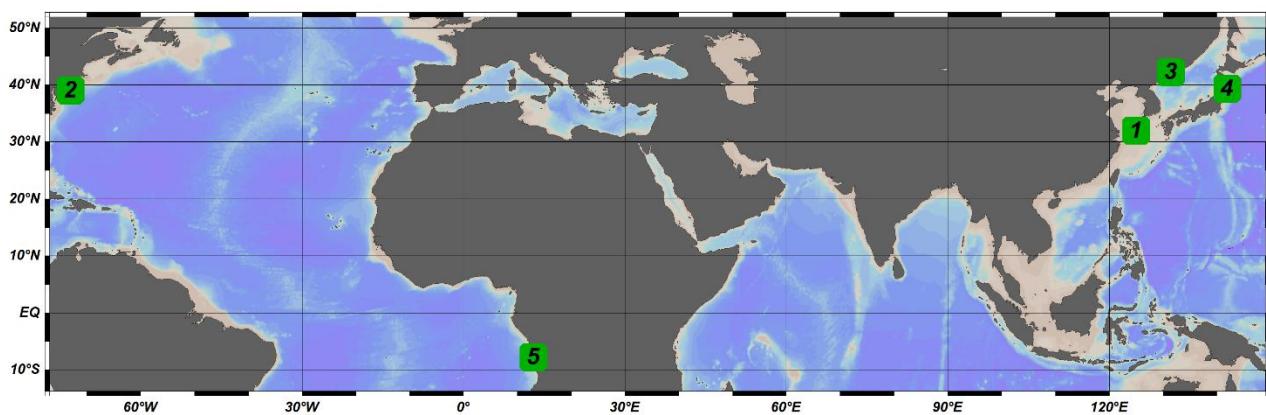


Рис. 1. Карта эпизодов наблюдений бризероподобных пакетов, описанных в литературных источниках

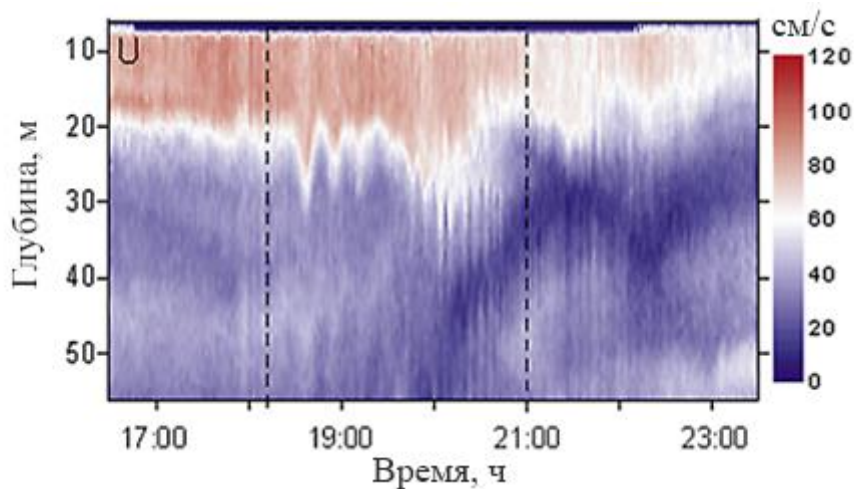


Рис. 2. Записи горизонтальных скоростей внутренних волн(см/с). Восточно-китайское море [Leeetal., 2006]

Распространение бризероподобных волн было зафиксировано на расстоянии 200км от побережья Южной Кореи [Leeetal., 2006], рис. 1. Наблюдение проводилось на протяжении 7 дней. Длительность волнового пакета составилаоколо часа, а высота – порядка 20 м. Число волн в пакете 4-5.

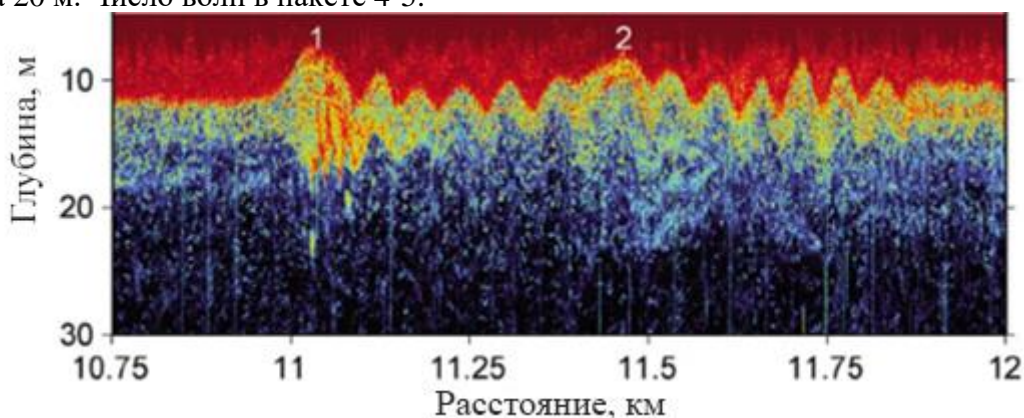


Рис. 3. Записи с акустического сенсора. Шельф Атлантического океана. [Shroyer et al., 2010]

11 августа 2006г. на шельфе Атлантического океана, в близи восточного побережья северной Америки были зафиксированы бризероподобные пакеты внутренних волн [Shroyer et al., 2010]. На рис. 3 видно несколько (по крайней мере два) таких пакетов. Их длина составляет около 300 м. Пакеты состоят из трех-четырех волн, максимальная высота которых достигает10м.

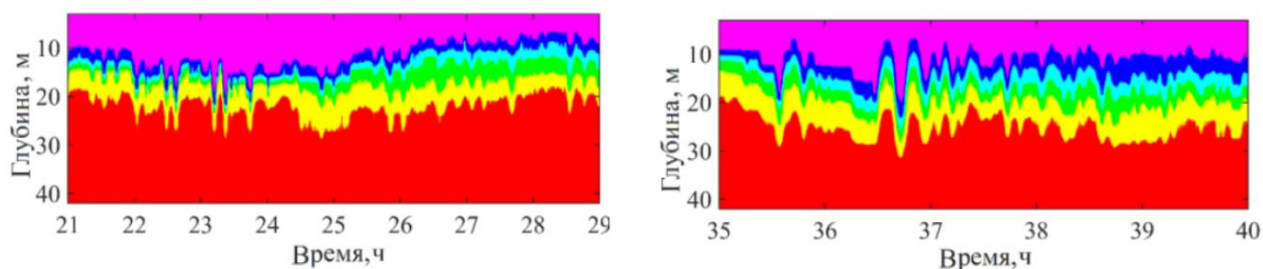


Рис. 4. Запись вертикального профиля температуры от времени. Японское море. [Кокоулина, 2020]

20-21 октября 2011г. были зафиксированы бризероподобные возмущения в заливе Петра Великого (Японское море) [Кокоулина, 2020], рис. 4. В отрезках записи на рис. 4 отчетливо видны не менее четырех локализованных волновых пакетов с длительностью от 20 до 50 минут и высотой 10-20 м, число волн в пакете 2-4.

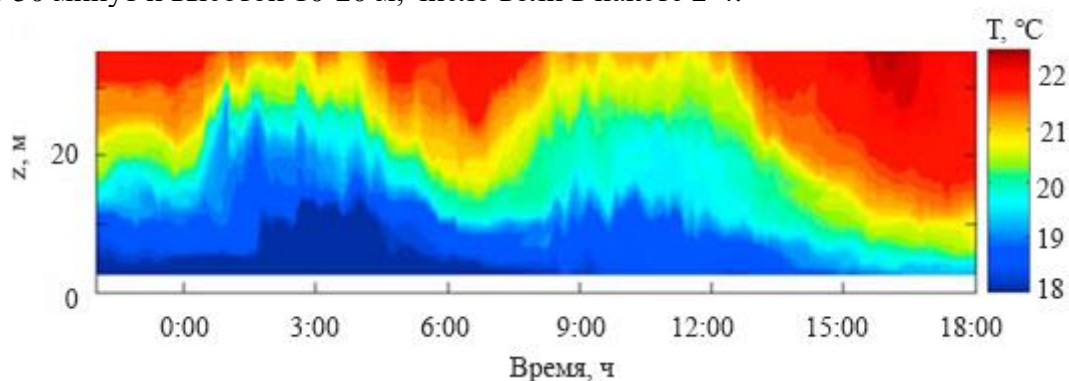


Рис. 5. Запись температуры в Японском море, бухта Оцучи. [Masunaga et al., 2016]

В сентябре 2013 года при исследовании, проводимом в бухте Оцучи, на востоке Японии [Masunaga et al., 2016] (рис. 5), в записи внутренних волн можно выделить бризероподобное возмущение длительностью около двух часов, содержащее три волны с максимальной высотой 9 м.

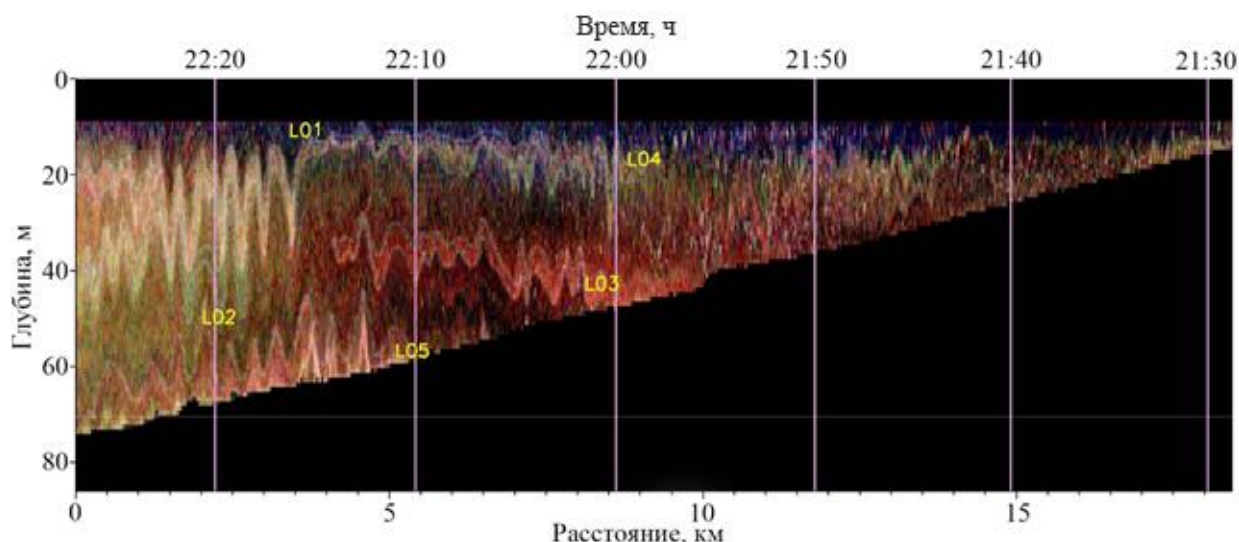


Рис. 6. Мульти-спектральная запись, побережье Амбриш [Ostrowski&Bazika-Sangolay, 2015]

В августе 2012г, на побережье города Амбриш, были зафиксированы бризероподобные волновые пакеты [Ostrowski & Bazika-Sangolay, 2015] (рис. 6). На рис. 6 можно наблюдать как минимум два нелинейных локализованных пакета. Длина пакетов

составляет 1-1,5км, которые состоят из двух - трех волн. Высота волнового пакета приблизительно 15 мс продолжительностью около пяти минут.

Все перечисленные выше эпизоды наблюдений объединены в таблицу (табл. 1), которая дает качественное представление о структуре локализованных волновых пакетов и их характерных особенностях.

Таблица 1

Сводная таблица наблюдений бризероподобных пакетов

№ эпизода	Расположение	Сезон	Кол-во волн	Высота волн	Длительность пакета/ длина	Источник
1	Восточно-Китайское море 32.00 С.Ш. 125.00 В.Д.	Лето (Август)	4-5	20 м	1ч	Lee et al., 2006
2	Шельф Атлантики 38.95 С.Ш. 72.93 З.Д.	Лето (Август)	3-4	10 м	300 м	Shroyer et al.,2010
3	Японское Море 42.27 С.Ш 131.5 В.Д	Осень (Октябрь)	2-4	10-20 м	20-50 мин	Кокоулина, 2020
4	БухтаОцучи 39.20С.Ш. 141.56 В.Д.	Осень (Сентябрь)	3	9 м	2 ч	Masunaga et al., 2016
5	Побережье Амбриш 7.51 Ю.Ш 13.5 В.Д	Лето (Август)	2-3	15 м	5 мин	Ostrowski & Bazika-Sangolay, 2015

Заключение

При анализе исследуемых эпизодов наблюдений внутренних волн (табл. 1) можно отметить, что бризероподобные пакеты наблюдались в самых различных акваториях Мирового океана в период с августа по октябрь. Практически во всех случаях общее количество наблюдаемых в пакете волн варьируется от двух до четырёх. Наиболее продолжительный пакет был зафиксирован в бухте Оцучи, его длительность составляет два часа. Самое короткое бризероподобное возмущение было зафиксировано на побережье Амбриш, его длительность порядка пяти минут, при этом высота достигает 15м. Необходимо отметить, что во всех рассмотренных эпизодах бризероподобные волны были зафиксированы при общей глубине акваторий до 100 м, так в первом эпизоде наблюдается волновой пакет высотой 20м, при общей глубине 84м. В дальнейшем планируется расширить коллекцию наблюдений, а также выявить принципиальные отличия и общие свойства локализованных волновых пакетов, встречающихся в природных условиях.

Представленные результаты получены при поддержке гранта Президента РФ по государственной поддержке ведущих научных школ РФ НШ-2485.2020.5 и молодых российских ученых – кандидатов наук МК-218.2020.5

Библиографический список

1. Кокоулина М. В., Куркина О. Е., Рувинская Е. А., & Куркин А. А. Вероятностные характеристики интенсивных короткопериодных внутренних волн в Японском море // Морской гидрофизический журнал. 2020. V.36. №5 P.215
2. Epifanova A.S., Rybin A.V., Moiseenko T.E., Kurkina O.E., Kurkin A.A., Tyugin D.Yu. Database of observations of the internal waves in the world ocean // Physical Oceanography. 2019. T. 26. № 4. С. 350-356.
3. Grimshaw R., Pelinovsky E, Talipova T. Solitary wave transformation in a medium with sign-variable quadratic nonlinearity and cubic nonlinearity // Physica D. 1999. V. 132. P. 40-62.
4. Lamb K., Polukhina O., Talipova T., Pelinovsky E., Xiao W., Kurkin A. Breather generation in the fully nonlinear models of a stratified fluid // Physical Rev. E. 2007. V. 75. No. 4. P. 046306.
5. Lee J.-H., Lozovatsky I., Jang S.-T., Jang Ch.-J., Hong Ch.-S., Fernando H.J.S. Episodes of nonlinear internal waves in the Northern East China Sea // Geophysical Research Letters. 2006. V. 33. P. L18601
6. Masunaga E., Fringer O. B., Yamazaki H., & Amakasu K. Strong turbulent mixing induced by internal bores interacting with internal tide-driven vertically sheared flow. Geophysical Research Letters. 2016. V. 43. № 5, P.2094–2101.
7. Ostrowski M., Bazika-Sangolay B. On physical mechanisms controlling inshore aggregations of small pelagic fish in a tropical upwelling system // 2015 IEEE/OES Acoustics in Underwater Geosciences Symposium (RIO Acoustics). IEEE. 2015. С. 1-7.
8. Rouvinskaya E., Talipova T., Kurkina O., Tyugin D., Soomere T. Transformation of internal breathers in the idealised shelf sea conditions // Continental Shelf Research. 2015. V. 110. P. 60-71
9. Shroyer E.L., Moum J.N., Nash J.D. Mode 2 waves on the continental shelf: Ephemeral components of the nonlinear internal wavefield // Jour. Geophys. Res. 2010a. V. 115. P. C07001.
10. Zhang X., Jie Z., Fan C., Wang, J. & Meng, J. Observations of internal waves with high sampling data of radar altimetry and MODIS images. // International Journal of Remote Sensing. 2018. V.39. № 21, P. 7405–7416.

L. Talalushkina, O. Kurkina, A. Giniyatullin

EPISODES OF OBSERVING BREATHLIKE INTERNAL WAVES IN THE WORLD OCEAN

Nizhny Novgorod state technical university n. a. R. E. Alekseev

The purpose of this work is to collect and systematize information on observations of breather-like internal wave packets in the World Ocean and to determine the water areas where the manifestation of the phenomena under study is theoretically possible. The work analyzed five records of internal waves available in the IGW Atlas database for the possible manifestation of breather-like wave packets in them.

Keywords: field data, internal waves, breather, three-layer fluid, database

ГИДРОЛОГИЧЕСКИЕ УСЛОВИЯ И ЭПИЗОДЫ НАБЛЮДЕНИЯ ВНУТРЕННИХ ВОЛН В БАЛТИЙСКОМ МОРЕ

Нижегородский государственный технический университет им. Р. Е. Алексеева

В работе проведен анализ вертикального изменения гидрологических характеристик, полученных в ходе экспедиции в Балтийском море в районе Готландской впадины. На основе исходных данных проведен анализ дисперсионных соотношений и структуры трёх низших бароклинных мод. Также в работе приведена оценка фазовой групповой скорости внутренних волн.

Ключевые слова: внутренние волны, дисперсионные зависимости, вертикальная мода, частота Вайсяля – Брента, Балтийское море.

Введение

Внутренние волны являются важным элементом динамической структуры океана и экосистем природных водоемов, они могут создавать значительные нагрузки и изгибающие моменты на подводные части гидротехнических сооружений, кроме того, существенно влияют на распределение питательных веществ [1-5]. Они наблюдаются в океане практически повсеместно, основной причиной их существования является устойчивая стратификация вод океана, нарушение равновесия которой приводит к возникновению волн на границах раздела слоев. Приливы являются основной, хотя и не единственной, причиной генерации внутренних волн. Как известно Балтийское море является бесприливным, но тем не менее в литературе имеется множество работ, показывающих наличие там внутренних волн. Кроме того, на сайте ИКИ РАН (http://iki.rssi.ru/asp/iw_images/index.html#bal-tic) имеются спутниковые снимки, подтверждающие этот факт. При моделировании внутренних волн необходимо знать гидрологические условия рассматриваемой акватории, так как они оказывают определяющее влияние на формирование различных режимов динамики волн.

Целью данной работы является исследование гидрологических условий существования внутренних волн в Балтийском море на основе натурных измерений, полученных в районе Готландской впадины.

Данные наблюдений

В работе рассмотрены натурные данные, полученные в летний период в районе Готландской впадины в Балтийском море. В ходе экспедиции были получены три записи длительностью несколько часов (от 4 до 8 часов). Эти данные включают в себя профили температуры, солёности, а также их колебания на верхнем и нижнем пикноклинах. Как показано в работе [6] в летние месяцы Балтийское море в большей степени является трехслойным, что подтверждает и стратификация в районе измерений (рис. 1.). Профиль плотности имеет два скачка, которые соответствуют пикам частоты Вайсяля–Брента. Резкое изменение температуры способствовало образованию верхнего пикноклина, галоклин же в свою очередь внес основной вклад в образование скачка плотности на глубине 60-80 м при общей глубине около 150 м. Замер профиля температуры, солёности проводился дважды: вначале измерений и в конце. Вертикальное распределение плотности было получено с помощью уравнения состояния воды TEOS-10 (<http://teos-10.org/>), встроенного в программный комплекс ODV (<https://odv.awi.de/>). Из рисунка 1 видно, что за 8 часов сильных изменений профиля плотности не произошло. Максимальное изменение температуры наблюдается на глубине около 30 метров и составляет 18% от начального значения, для солёности – примерно 15% на глубине 70 метров.

На восьмичасовой записи смещения плотности на верхнем пикноклине, глубина залегания которого составляет примерно 25 м, наблюдаются волны амплитудой до 6 м. За первые 1.5 часа записи смещения изопикны, которая расположена на глубине примерно 65 м, наблюдаются волны с амплитудами, максимум которых не превышает 4-5 м, и по форме напоминающие солитоноподобные возмущения, длительностью 5-10 минут.

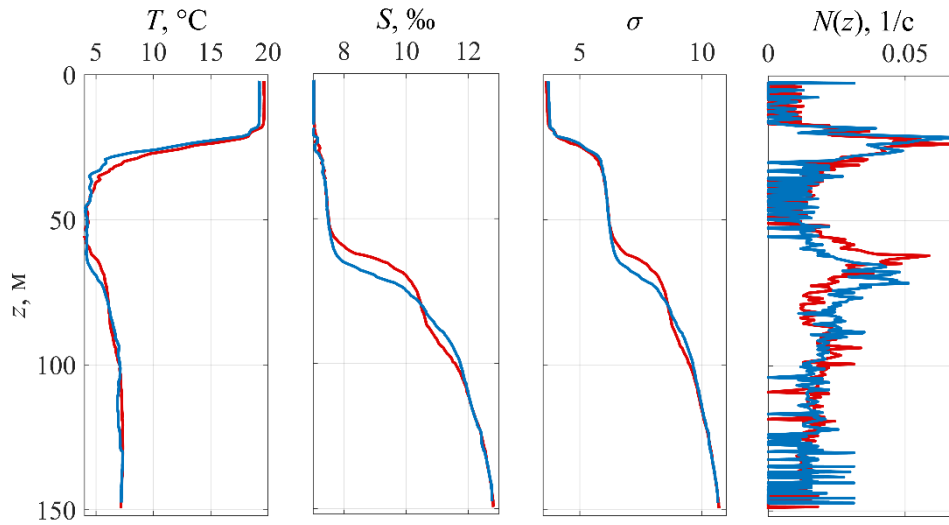


Рис. 1. Вертикальный профиль температуры, солёности, условной плотности и частоты Вайсяля – Брента, синим цветом, представлены данные, полученные в начале периода измерений, красным – в конце.

Исследование вертикальной структуры поля внутренних волн заключается в решении задачи Штурма – Лиувилля следующего вида:

$$\frac{d^2\Phi}{dz^2} + k^2 \frac{N^2(z) - \omega^2}{\omega^2 - f^2} \Phi = 0 \quad (1)$$

с граничными условиями $\Phi(z=0) = \Phi(H=0) = 0$, где z – вертикальная координата, направленная вниз, $z = 0$ – невозмущенная поверхность моря; H – полная глубина бассейна в расчетной точке; $\Phi(z)$ – вертикальная структура внутренней волны; k – волновое число; ω – частота внутренних волн, $f = 2\Omega \sin \varphi$ – инерционная частота (Ω – частота вращения Земли, соответствующая периоду 24 ч, φ – географическая широта точки наблюдений), $N(z) = \sqrt{\frac{g}{\rho(z)} \frac{d\rho(z)}{dz}}$ – частота Вайсяля-Брента (плавучести), g – ускорение свободного падения, $\rho(z)$ – вертикальный профиль плотности морской воды. Для удобства записи и расчетов мы используем условную плотность $\sigma = \rho(z) - 1000 \text{ кг/м}^3$. Стоит также отметить, что частоты внутренних волн лежат в диапазоне $f \leq \omega \leq N_{\max}$, N_{\max} – максимум $N(z)$, для данного района измерений он составляет примерно 0.06 1/с.

Известно, что решение задачи (1) содержит бесконечный набор собственных функций – мод $\Phi_n(k, z)$ и собственных значений $\omega_n(k)$, в данной работе мы остановились на рассмотрении трех низших мод ($n=1,2,3$). Для длинных волн решение этой задачи представлено в работах [7-13]. Для каждого значения волнового числа k , которое менялось в диапазоне от 0 до $k_{\max} = 2\pi/H$ шагом $k_{\max}/100$, численно (методом Рунге-Кутты) решается дифференциальное уравнение (1) с начальными условиями $\Phi(0) = 0$, $\Phi'(0) = \Phi_0$ методом «стрельбы». Метод заключается в варьировании пристрелочного параметра до тех пор, пока для решения не будут с заданной точностью выполняться граничные условия. В данной однородной задаче функция $\Phi(z)$ определяется с точностью до множителя-константы, поэтому пристрелочным параметром является не Φ_0 , как в стандартном

варианте метода, либо волновое число k либо частота ω , входящие в множитель перед $\Phi(z)$ в уравнение (1). В качестве пристрелочного параметра выбрана частота внутренних

волн, с начальным значением $\omega = \sqrt{\frac{(N_{\max} kH)^2 + (\pi n f)^2}{(kH)^2 + (\pi n)^2}}$, где n – номер моды. Поскольку

каждая мода $\Phi_n(z)$ определяется с точностью до множителя-константы, его удобно выбрать так, чтобы первый экстремум этой функции на отрезке $z \in [0; H]$ был равен единице. Это обеспечивает однозначность в определении вертикальных мод, необходимую, например, при сравнении решений в разных точках бассейна. На рис. 2 показана структура первых трёх модовых функций для рассматриваемой акватории, рассчитанных для стратификации в момент начала измерений. Стоит отметить, что с увеличением волнового числа, ширина зоны максимума первой моды сужается.

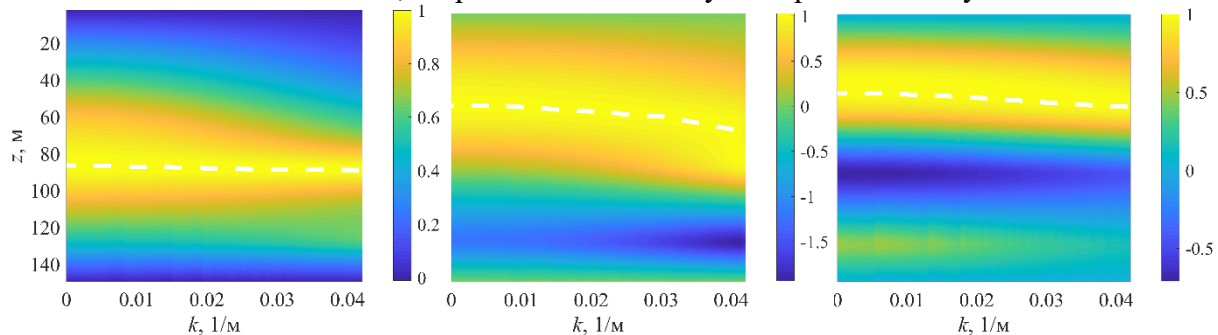


Рис 2. Первые три низшие нормализованные вертикальные моды в зависимости от волнового числа k . Цветом показано значение модовой функции $\Phi_n(k, z)$, $n=1,2,3$ слева направо, соответственно.

Из рис. 3, на котором слева приведены дисперсионные кривые $\omega_n(k)$, видно, что для рассмотренной акватории Балтийского моря присуще плавное изменение частоты внутренних волн с уменьшением длины волны, а также отсутствуют зоны сгущения и резонанса в диапазоне допустимых волновых чисел k и частот ω [14, 15].

Из дисперсионного соотношения не сложно вычислить групповую (c_g) и фазовую скорости (c_p), по формулам:

$$c_p = \frac{\omega}{k}, \quad c_g = \frac{d\omega}{dk}$$

На рис. 3 справа представлен график изменения групповой и фазовой скорости волн.

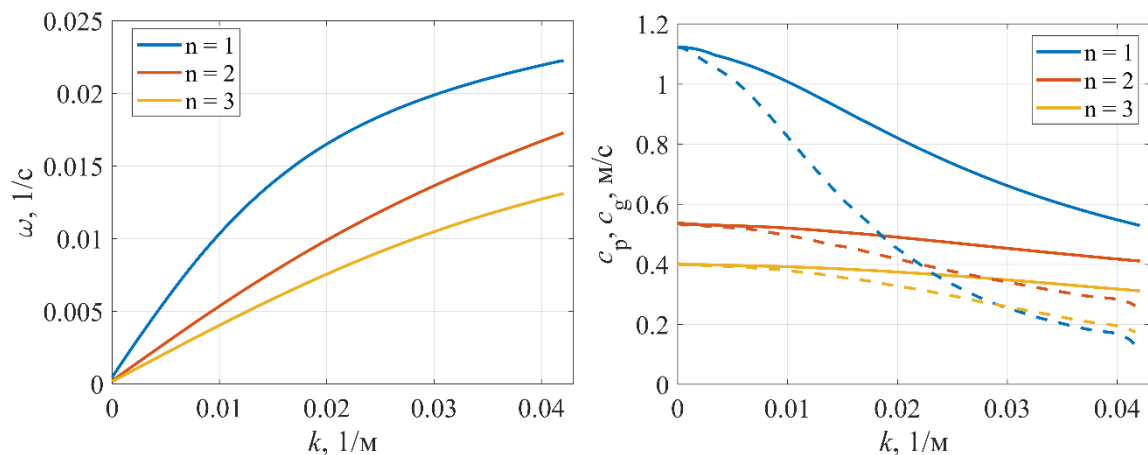


Рис. 3. Слева: дисперсионные кривые трех низших мод внутренних волн, справа: сплошные линии – фазовая скорость c_p , пунктирные – групповая скорость c_g в зависимости от волнового числа

В контексте линейной негидростатической модели [16] неоднородный по времени поток над неровным дном вызывает резонансные бароклинные движения с частотой ω ,

когда выполняется условие $\alpha \leq \gamma$ для наклона α характеристик гиперболического волнового уравнения:

$$\alpha^2(z) = \frac{\omega^2 - f^2}{N^2(z) - \omega^2}$$

и уклона дна $\gamma = dH/dx$, где $H(x)$ – глубина дна в зависимости от горизонтальной координаты. Эти надкритические области, где движения жидкости являются существенно бароклинными во всей толще, являются источниками излучения волн создавая волновые лучи (beams), определяемые из уравнения:

$$\frac{dz}{dx} = \pm \sqrt{\frac{\omega^2 - f^2}{N^2(z) - \omega^2}},$$

решение которого записывается в виде квадратур:

$$x - x_0 = \int_{z_0}^z \frac{dz}{\sqrt{\frac{\omega^2 - f^2}{N^2(z) - \omega^2}}}.$$

Распространение внутренних волновых пучков широко обсуждалось в литературе [17 - 19]. Эти лучи (или пучки) часто заметны в расчетах внутренних волновых полей в рамках полнонелинейных негидростатических моделей. На рисунке 4 представлен рассчитанный численно луч, для исследуемой акватории, при условии отсутствия неровности дна, вдоль которого распространяются внутренние волны с отражением от морской поверхности и от дна. Для известной батиметрии бассейна лучи могут быть рассчитаны с учетом неоднородностей дна.

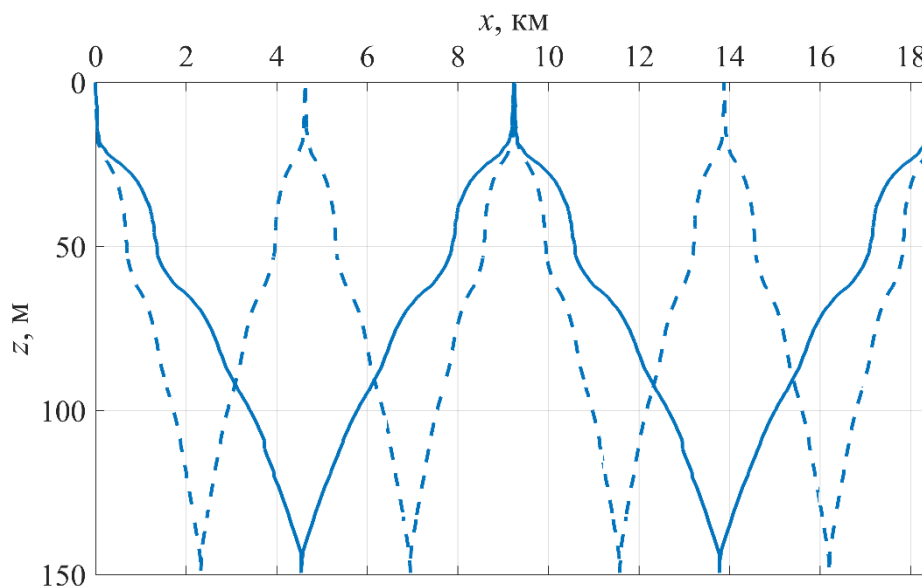


Рис. 4. Структура внутренних волновых лучей частотой, соответствующей периоду один час (сплошная линия) и соответствующая периоду 30 мин (штриховая линия) для исследуемой акватории

Заключение

В работе проведен анализ натуральных данных, полученных в Балтийском море в районе Готландской котловины. Показано, что вертикальная стратификация плотности в исследуемой акватории является квазитрехслойной. На основе данных прямых контактных измерений изучена модовая структура поля внутренних волн, их дисперсионные свойства, проведена оценка фазовой и групповой скорости. Построены траектории лучей распространения внутренних волн. Выполненные расчеты пополняют океанографическую геоинформационную онлайн систему.

Представленные результаты получены при поддержке гранта Президента РФ по государственной поддержке ведущих научных школ РФ НШ-2485.2020.5 и стипендии Президента Российской Федерации молодым ученым и аспирантам (СП-692.2021.5).

Библиографический список

1. Cai S. A method to estimate the forces exerted by internal solitons on cylindrical piles / S. Cai, X. Long, Z. Gan // *Ocean Engineering*. – 2003. – Т. 30. – №. 5. – С. 673-689.
2. Cai S. Forces and torques exerted by internal solitons in shear flows on cylindrical piles / S. Cai, X. Long, S. Wang // *Applied Ocean Research*. – 2008. – V. 30. – № 1. – P. 72-77.
3. Chakrabarti S. *Handbook of Offshore Engineering (2-volume set)* / S. Chakrabarti // Elsevier. – 2005.
4. Osborne A. Nonlinear ocean wave and the inverse scattering transform / A. Osborne // *Scattering*. – Academic Press, 2002. – P. 637-666.
5. Woodson C. The fate and impact of internal waves in nearshore ecosystems / C. Woodson // *Annual review of marine science*. – 2018. – V. 10. – P. 421-441.5.
6. Рувинская Е.А. Зонирование по типам плотностной стратификации вод Балтийского моря в контексте динамики внутренних гравитационных волн / Е. А. Рувинская, Д. Ю. Тюгин, О. Е. Куркина, А. А. Куркин // *Фундаментальная и прикладная гидрофизика*. – 2018. – Т. 11. – №. 1. – С. 46-51.
7. Pelinovsky E.N., Internal Solitary Waves. In Book: *Solitary Waves in Fluids* (ed. By R. Grimshaw) / E. N. Pelinovsky, A. V. Slunyaev, O. E. Polukhina, T. G. Talipova // WIT Press, Southampton, Boston. – 2007. – С. 85-11.
8. Kurkina O. Mapping the internal wave field in the Baltic Sea in the context of sediment transport in shallow water / O. Kurkina, T. Talipova, E. Pelinovsky, T. Soomere // *Journal of Coastal Research*. – 2011. – С. 2042-2047.
9. Kurkina O. Propagation regimes and populations of internal waves in the Mediterranean Sea basin / O. Kurkina, E. Rouvinskaya, T. Talipova, T. Soomere // *Estuarine, Coastal and Shelf Science*. – 2017. – Т. 185. – С. 44-54.
10. Kurkina O. Pycnocline variations in the Baltic Sea affect background conditions for internal waves / O. Kurkina, A. Kurkin, T. Soomere, A. Rybin, D. Tyugin // *2014 IEEE/OES Baltic International Symposium (BALTIC)*. – IEEE, 2014. – С. 1-8.
11. Kurkina O. Kinematic parameters of internal waves of the second mode in the South China Sea / O. Kurkina, T. Talipova, T. Soomere, A. Giniyatullin, A. Kurkin // *Nonlinear Processes in Geophysics*. – 2017a. – Т. 24. – №. 4.
12. Kurkin A. Internal waves in the Mediterranean Sea: mapping and transformations / A. Kurkin, O. Kurkina, T. Talipova, E. Pelinovsky, E. Rouvinskaya // *12th International Conference on the Mediterranean Coastal Environment, MEDCOAST 2015*. – 2015. – С. 835-846.
13. Pelinovsky E. N. Modelling of internal waves in the Baltic Sea / E. N. Pelinovsky, T. G. Talipova, T. Soomere, O. E. Kurkina, A. A. Kurkin, D. Y. Tyugin // *Фундаментальная и прикладная гидрофизика*. – 2018. – Т. 11. – №. 2. – С. 8-20.
14. Grigorenko K. S. Internal Waves Kinematic Characteristics in the Central Atlantic in 2014–2015 / K. S. Grigorenko, S.M. Khartiev // *Physical and Mathematical Modeling of Earth and Environment Processes (2018)*. – 2019. – P. 22-32.
15. Grigorenko K. S. Investigations of Internal Waves in the Seas of Russia and in the Central Atlantic / K. S. Grigorenko, S. M. Khartiev // *Conference on Physical and Mathematical Modeling of Earth and Environment Processes*. – 2017. – P. 28-34.
16. Baines D. *Emigration from Europe 1815-1930* / D. Baines // Cambridge University Press. – 1995. – №. 11.

17. Kurkina O. E. Numerical Modeling of Internal Wave Generation at High Latitudes /O. E. Kurkina, T. G. Talipova, E. N. Pelinovsky, A. A. Kurkin//The Ocean in Motion. – Springer, Cham. –2018. – C. 569-580.
18. Gerkema T. Internal tides in the Bay of Biscay: conversion rates and seasonal effects /T. Gerkema, F. P. A. Lam, L. R. M. Maas//Deep Sea Research Part II: Topical Studies in Oceanography. – 2004. – T. 51. – №. 25-26. – C. 2995-3008.
19. Grimshaw R. Nonreflecting internal wave beam propagation in the deep ocean / R. Grimshaw, E. Pelinovsky, T. Talipova //Journal of physical oceanography. – 2010. – T. 40. – №. 4. – C. 802-813.

M. Kokoulina, O. Kurkina, A. Kurkin, L. Kuznetsova

**HYDROLOGICAL CONDITIONS AND EPISODES OF OBSERVATION OF
INTERNAL WAVES IN THE BALTIC SEA**

Nizhny Novgorod state technical university n. a. R. E. Alekseev

The paper analyzes the vertical changes in hydrological characteristics obtained during an expedition to the Baltic Sea in the Gotland Basin. Based on the initial data, the analysis of dispersion relations and structure of three lower baroclinic modes is carried out. The paper also provides an estimate of the phase and group velocities of internal waves.

Keywords: internal waves, dispersion dependences, vertical mode, Väisälä - Brent frequency, Baltic Sea.

ЧИСЛЕННОЕ МОДЕЛИРОВАНИЕ РАСПРОСТРАНЕНИЯ ИНТЕНСИВНЫХ ВНУТРЕННИХ ВОЛН В ДВУХСЛОЙНОЙ ЖИДКОСТИ

*Нижегородский государственный технический университет им Р.Е. Алексеева¹
Российский федеральный ядерный центр - Всероссийский научно-исследовательский
институт экспериментальной физики, Саров²*

Целью данной работы является построение трехмерного цифрового двойника экспериментальной гидродинамической установки в НИЛ МПиТК НГТУ им. Р.Е. Алексеева, построение сетки для численного решения и осуществление численного моделирования динамики двухслойной жидкости в лотке при помощи программного обеспечения ЛОГОС.

Ключевые слова: математическое моделирование, вычислительная гидродинамика, стратифицированная жидкость, двухслойная жидкость, внутренние волны, программный пакет ЛОГОС.

Введение

Исследование внутренних волн, в особенности их влияния на окружающую среду, является одним из актуальных направлений развития современной гидродинамики.

Для численного исследования внутренних волн в стратифицированных жидкостях используются различные программные обеспечения, такие как: глобальная модель циркуляции MITgcm [1], ЛОГОСАэро-Гидро [2], численная модель ISWfoam, вычислительный комплекс IGWResearch [3] и другие программные решения.

В данной работе для математического моделирования динамики внутренних волн в двухслойной жидкости была выбрана для использования программа ЛОГОС Аэро-Гидро – одна из нескольких в пакете программ ЛОГОС, разработкой которого уже много лет занимается РФЯЦ-ВНИИЭФ в городе Саров.

На текущий момент пакет программ ЛОГОС имеет широкий спектр применения, в частности он использовался для оценки гидродинамических характеристик корпуса и судовых движителей, расчета обтекания обводов и оценки возможности снижения гидродинамического сопротивления корпуса на этапе проектирования [4], для выявления эффектов, которые сопровождают вхождение астероида в водную среду [5] и во многих других экспериментах [6-9]. Полученные в ходе численного моделирования результаты можно применять как в рамках производства: введения инновационных технологий, так и в совершенствовании повседневной жизни человека.

В данной работе описывается процесс создания цифрового двойника гидродинамической установки НИЛ МПиТК, построение его вычислительной сетки, а также визуализация динамики двухслойной жидкости.

Численное моделирование динамики жидкости в пакете программ ЛОГОС

Для моделирования динамики несжимаемой жидкости в ЛОГОС используется математическая модель многофазной среды, основанная на системе уравнений Навье-Стокса. Предполагается, что течение изотермическое, а поле скорости общее для всех фаз. С учётом подобных допущений получим систему следующих уравнений: уравнение сохранения массы, уравнение сохранения импульса и уравнение переноса объёмной доли, которая в декартовых координатах выглядит таким образом:

$$\begin{cases} \frac{\partial u_i}{\partial x_i} = 0, \\ \frac{\partial u_i}{\partial \tau} + \rho \frac{\partial}{\partial x_j} (u_i u_j) = -\frac{\partial p}{\partial x_i} + \frac{\partial}{\partial x_j} \tau_{ij} + \rho g_i, \\ \frac{\partial \alpha_w}{\partial \tau} + \frac{\partial}{\partial x_i} (u_i \alpha_w) = 0, \end{cases} \quad (1)$$

где i, j – нижние индексы, указывающие на принадлежность векторных координат декартовым, $i, j = \{x, y, z\}$, ρ – средняя плотность, вычисляемая по формуле: $\rho = (\rho_w \alpha_w + \rho_{sw} \alpha_{sw})$, w – индекс принадлежности к фазе «пресная вода», sw – индекс принадлежности к фазе «солёная вода», α_w – объемная доля пресной воды, u_i – компонента вектора скорости, $i = \{x, y, z\}$, t – время, p – давление, x_i – компонента вектора в декартовых координатах, $i = \{x, y, z\}$, τ_{ij} – тензор вязких напряжений, который по гипотезе Буссинеска имеет вид: $\tau_{ij} = \mu \left(\frac{\partial u_i}{\partial x_j} + \frac{\partial u_j}{\partial x_i} + \frac{2}{3} \frac{\partial u_k}{\partial x_k} \delta_{ij} \right)$, где μ – динамическая вязкость, δ_{ij} – символы Кронекера, g_i – компонента вектора свободного падения [10].

Для решения системы (1) в пакете программ ЛОГОС реализован численный метод VOF (Volumeoffluid) [11-13].

Сила тяжести учитывается с использованием алгоритма, который основан на поправке объёмных сил [14], что обеспечивает отсутствие паразитных осцилляций на сетках любого типа.

Условия задачи

На базе НИЛ МПиТК установлен гидродинамический лоток, который позволяет проводить различные эксперименты по динамике двухслойной жидкости. Лабораторный бассейн представляет собой прямоугольный резервуар из органического стекла, размеры которого указаны в Табл. 1. Лабораторная установка оснащена вертикальной перегородкой, которая позволяет герметично разделять различные уровни жидкости. Один из вариантов начальной задачи о гидродинамическом коллапсе в двухслойной среде, реализованная в виде серии лабораторных экспериментов, представлена на Рис. 1.

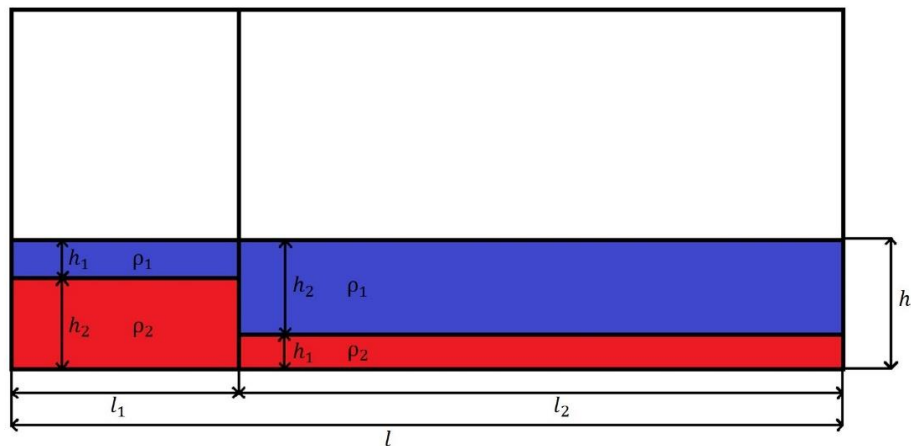


Рис. 1. Схема задачи.

Геометрия задачи такова: $l_1=0.45\text{м}$ – расстояние от левой стенки установки до перегородки между отсеками, $l_2=5.91\text{м}$ – расстояние от перегородки между отсеками до правой стенки установки, $h_1 = 0.05\text{м}$ – толщина слоя пресной воды в левом отсеке и толщина слоя солёной воды в правом отсеке, $h_2=0.2\text{м}$ – толщина слоя солёной воды в левом отсеке и толщина слоя пресной воды в правом отсеке, $\rho_1 = 995.75 \text{ кг/м}^3$ – плотность пресной воды, $\rho_2 = 1019.49 \text{ кг/м}^3$ – плотность солёной воды. Параметр динамической вязкости μ в системе уравнений (1) принимался равным $1 \cdot 10^{-6} \text{ кг/(м}\cdot\text{с)}$.

Параметры расчетной области для численных экспериментов

	Физические размеры (метры)	Размеры рабочей области (метры)	Сетка (точки разбиения)
Длина (ось X)	6.36	6.36	1000
Высота (ось Y)	1	0.25	50
Ширина (ось Z)	0.5	0.5	1

Расчётная сетка

В настоящей работе проводится численное моделирование динамики интенсивных внутренних волн в двухслойной жидкости: решается задача Коши в рамках уравнений (1) с исходными условиями, аналогичными условиям в лабораторных экспериментах, параметры которых приведены выше.

Расчётная область в рамках модели конфигурировалась так же, как и реализованная лабораторная ситуация, а именно, рассматривалась двумерная область в вертикальной плоскости (Табл. 1), заполненная двухслойной жидкостью. В результате численного моделирования был создан цифровой двойник рабочей области гидродинамической установки.

После построения трехмерной реализации, была построена объемная расчётная сетка. Так как движение в рамках данных условий происходит преимущественно квази-двумерно, вдоль лотка (по плоскости OXY), то движением по оси Z можно пренебречь и выбрать на ней минимально возможное число точек разбиения расчётной сетки. Таким образом, была выбрана сетка со следующим количеством точек разбиения по каждому из измерений, представленных в Табл. 1.

В соответствии с заданными размерами сетки, размер ячейки по оси X равен 6.36 мм, размер ячейки по оси Y равен 5 мм.

Расчётная сетка является важным моментом в любой численной задаче. Как правило, чем больше точек – тем точнее численное решение. Однако, стоит понимать, что большое число точек требует высокой производительности системы. Если система не будет высокопроизводительной, то решение численной задачи будет более затратно по времени. Для точек разбиения в цифровом двойнике экспериментальной гидродинамической установки и по длине, и по высоте было выбрано равномерное распределение, т.е. точки находятся на равном расстоянии друг от друга.

Так как сетка имеет большое число узлов, то вычисления происходили длительное время. На компьютере со следующими характеристиками вычисления заняли ≈ 10 часов: процессор Intel Core i7-6700k, видеокарта NVIDIA GTX 1070, ОЗУ Kingston KHX2400c15/8g DDR4-2400 (2 шт., 8гб каждая).

Результаты

В процессе эксперимента удаление вертикальной перегородки приводит к тому, что тяжелая жидкость, которая находится в левой части бассейна, начинает перетекать в правую часть, вместе с чем, происходит перемещение легкой жидкости из правой части в левую область бассейна. Вследствие этого происходит образование внутренней уединенной волны положительной полярности и последующее ее распространение (Рис. 2).

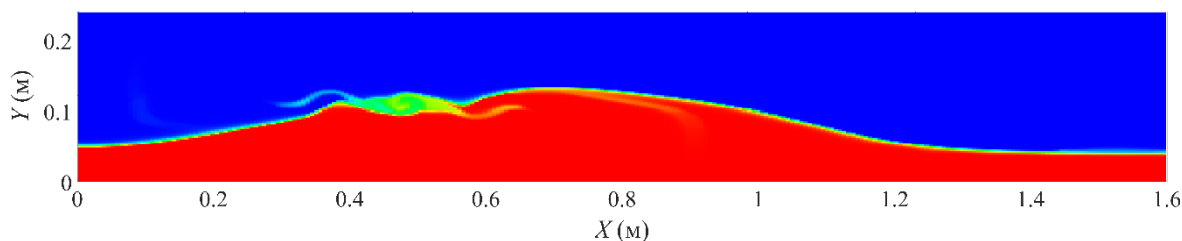


Рис. 2. Процесс формирования уединённой волны в численном эксперименте.

На рис. 3 проиллюстрирован момент взаимодействия внутренней уединенной волны положительной полярности с правой границей расчетной области (правая стенка лотка). Необходимо отметить, что при взаимодействии волны со стенкой ее амплитуда увеличилась практически в 1.5 раза. Детальное описание процесса взаимодействия волны с преградой важно для расчетов величин волнового давления и динамических сил, вызываемых пульсацией скоростей индуцированного волной потока.

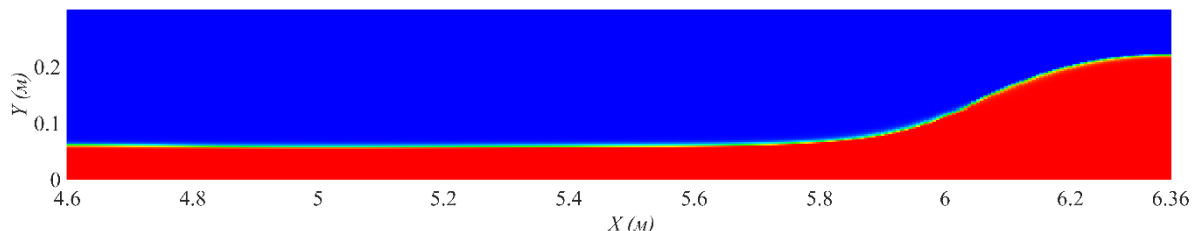


Рис. 3. Взаимодействие уединенной волны со стенкой гидродинамической установки.

Заключение

При помощи программного пакета ЛОГОС был построен трехмерный цифровой двойник экспериментальной гидродинамической установки в НИЛ МПиТК НГТУ им. Р.Е. Алексеева. Был смоделирован процесс генерации внутренних волн методом гидродинамического коллапса в двухслойной жидкости и ее дальнейшее распространение.

В перспективе планируется проведение серии численных экспериментов с варьированием параметров стратификации жидкости, и продолжение работ по сравнению волновых полей с полученными в рамках лабораторных экспериментов. Целью этого является калибровка численной модели и доведение процесса расчетов до уровня, когда возможно будет заменить большую часть физических экспериментов численными.

Представленные результаты получены в рамках государственного задания в сфере научной деятельности № FSWE -2020-0007.

Библиографический список

1. Adcroft A., Campin J.-M. MITgcm User Manual. Cambridge: MIT. 2011. 455 p.
2. Дерюгин Ю. Н., Козелков А. С., Зеленский Д. К., Глазунов В. А., Полищук С. Н., Ялозо А. В., Емельянов А. Б. Программный модуль пакета программ "ЛОГОС" для моделирования аэрогидродинамики и теплопереноса (ЛОГОС-ТМП), версия 2017. – 2019.
3. Lamb K. G. Numerical experiments of internal wave generation by strong tidal flow across a finite-amplitude bank edge // J. Geophys. Res. Oceans. 1994. V. 99. P. 843-864.
4. Применение Логос в судостроении // РФЯЦ-ВНИИЭФ – URL: <http://logos.vniief.ru/papPLIC/shipbuilding/8011fe004de4da14b9dabd5d660404d1>
5. Козелков А. С. Эффекты, сопровождающие вхождение астероида в водную среду // Труды НГТУ им. Р.Е. Алексеева. 2014. №. 3 (105).
6. Применение Логос в энергетике // РФЯЦ-ВНИИЭФ – URL: <http://logos.vniief.ru/papPLIC/energy/8011fe004de4da14b9dabd5d660404d1>
7. Применение Логос в автомобилестроении// РФЯЦ-ВНИИЭФ – URL: <http://logos.vniief.ru/papPLIC/transport/8011fe004de4da14b9dabd5d660404d1>
8. Ракетно-космическая отрасль// РФЯЦ-ВНИИЭФ – URL: <http://logos.vniief.ru/papPLIC/space/8011fe004de4da14b9dabd5d660404d1>
9. Применение Логос в авиастроении // РФЯЦ-ВНИИЭФ – URL: <http://logos.vniief.ru/papPLIC/aviation/>
10. Плыгунова К. С., Тятюшкина Е. С., Козелков А. С., Курулин В. В., Уткин Д. А. Моделирование плавания тел на базе метода VOF совместно с методом деформации сетки // КОГРАФ-2020. С. 324-329.

11. Hirt C.W., Nichols B.D. Volume of fluid (VOF) method for the dynamics of free boundaries // *J. Comput. Phys.*, Vol. 39, p. 201-225, 1981
12. Deryugin Yu., Zhuchkov R., Zelenskiy D., Kozelkov A., Sarazov A., Kudimov N., Lipnickiy Yu., Panasenko A. Validation results for the LOGOS multifunction software package in solving problems of aerodynamics and gas dynamics for the lift-off and injection of launch vehicles // *Mathematical Models and Computer Simulations*. 2015, V.7. Issue 2. P. 144–153
13. Курулин В.В., Козелков А.С., Ефремов В.Р., Яцевич С.В., Тарасова Н.В. Применение метода VOF для решения комплексных задач со свободной поверхностью // XXV Всероссийский семинар с международным участием по струйным, отрывным и нестационарным течениям, 2018, С. 144-145
14. Ефремов В.Р., Козелков А.С., Корнев А.В., Куркин А.А., Курулин В.В., Стрелец Д.Ю., Тарасова, Н.В. Метод учета сил гравитации при моделировании течений со свободной поверхностью // *Журнал вычислительной математики и математической физики*. 2017. Т. 57. № 10

M. E. Shokhov¹, L. V. Talalushkina¹, V. V. Kurulin^{1,2}, A. A. Kurkin¹, O. E. Kurkina¹

NUMERICAL SIMULATION OF THE PROPAGATION OF INTENSE INTERNAL WAVES IN A TWO-LAYER LIQUID

*Nizhny Novgorod State Technical University named after R. E. Alekseev
Nizhny Novgorod¹, Russian Federal Nuclear Center - All-Russian Research Institute of
Experimental Physics
Sarov²*

The purpose of this work is to construct a three-dimensional digital double of an experimental hydrodynamic installation in the LMNAD NSTU named after R. E. Alekseev, to construct a grid for numerical solution and to perform numerical simulation of the dynamics of a two-layer liquid in a tray using the LOGOS software.

Keywords: mathematical modeling, computational fluid dynamics, stratified fluid, two-layer fluid, internal waves, LOGOS software package.

О МЕРЕ ЛЕБЕГА АТТРАКТОРОВ, ЗАДАННЫХ ГОМОТЕТИЯМИ С АФФИННО НЕЗАВИСИМЫМИ ЦЕНТРАМИ

Нижегородский государственный технический университет им. Р.Е. Алексеева

Рассматриваются аттракторы систем итерированных функций в n -мерном евклидовом пространстве, состоящих из $n + 1$ гомотетии с произвольными коэффициентами подобия и аффинно независимыми центрами. Найдены условия, при которых мера Лебега таких аттракторов равна нулю или положительна.

Ключевые слова: система итерированных функций, аттрактор, фрактал, мера Лебега, множество Кантора

Введение

Конечное семейство $S = \{f_i, i = 1 \dots n\}$ сжимающих отображений, заданных на полном метрическом пространстве X , называется системой итерированных функций (СИФ). Согласно известной теореме Хатчинсона для любой СИФ существует единственное непустое компактное подмножество $A \subset X$, инвариантное относительно S :

$$f_1(A) \cup f_2(A) \cup \dots \cup f_n(A) = A.$$

Компактное множество A называется аттрактором СИФ S .

Пусть $S = \{f_i | i = 1..n + 1\}$ – система итерированных функций в евклидовом пространстве R^n , состоящая из $n + 1$ гомотетии с коэффициентами подобия $\lambda_i \in (0,1)$ и аффинно независимыми центрами A_1, A_2, \dots, A_{n+1} . Напомним, что точки A_1, A_2, \dots, A_{n+1} в R^n называются аффинно независимыми, если они не лежат в подпространстве размерности $n - 1$. Если $n = 2$ и $\lambda_1 = \lambda_2 = \lambda_3 = 1/2$, то аттрактором такой СИФ является треугольник (салфетка) Серпинского.

Аттракторы СИФ S исследовались в [1-3], а также авторами в [4,5].

В [1] предполагается, что все коэффициенты гомотетии равны: $\lambda_1 = \lambda_2 = \dots = \lambda_{n+1} = \lambda$. Показано [1, Предложение 4.3], что при $\lambda < (n + 1)^{-1/n}$ n -мерная мера Лебега аттрактора СИФ S равна нулю. Если $\lambda \geq n/(n + 1)$, то аттрактор A представляет собой симплекс Δ с вершинами – центрами гомотетий; следовательно, мера Лебега аттрактора A совпадает с объемом симплекса Δ .

Цель данной работы – распространить результаты Бромхеда, Монталди, Сидорова [1] о мере Лебега, указанные выше, на случай аттракторов, заданных системами итерированных гомотетий с неравными коэффициентами подобия.

Основные утверждения и их доказательство

Теорема 1. Пусть $S = \{f_i | i = 1..n + 1\}$ – система итерированных функций в евклидовом пространстве R^n , состоящая из $n + 1$ гомотетии с коэффициентами подобия $\lambda_i \in (0,1)$ и аффинно независимыми центрами, A – аттрактор СИФ S . Если

$$\lambda_1^n + \lambda_2^n + \dots + \lambda_{n+1}^n < 1, \tag{1}$$

то n -мерная мера Лебега аттрактора A равна нулю.

Доказательство. Обозначим через μ n -мерную меру Лебега в R^n . Поскольку f_i – аффинное преобразование n -мерного пространства R^n с коэффициентом сжатия λ_i , то $\mu(f_i(A)) = \lambda_i^n \mu(A), i = 1..n + 1$.

Так как A – аттрактор СИФ S , то $A = f_1(A) \cup f_2(A) \cup \dots \cup f_{n+1}(A)$. В силу полуаддитивности меры Лебега имеем

$$\begin{aligned} \mu(A) &= \mu(f_1(A) \cup f_2(A) \cup \dots \cup f_{n+1}(A)) \leq \\ &\leq \mu(f_1(A)) + \mu(f_2(A)) + \dots + \mu(f_{n+1}(A)) = \\ &= \lambda_1^n \mu(A) + \lambda_2^n \mu(A) + \dots + \lambda_{n+1}^n \mu(A) = \\ &= (\lambda_1^n + \lambda_2^n + \dots + \lambda_{n+1}^n) \mu(A). \end{aligned}$$

Таким образом,

$$\mu(A) \leq (\lambda_1^n + \lambda_2^n + \dots + \lambda_{n+1}^n) \mu(A).$$

Если $\mu(A) \neq 0$, то $\lambda_1^n + \lambda_2^n + \dots + \lambda_{n+1}^n \geq 1$, что противоречит условиям теоремы. Следовательно, $\mu(A) = 0$.

В случае равенства коэффициентов подобия получаем результат *Бромхеда, Монталди и Сидорова [1], упомянутый выше*.

Следствие 1. Если $\lambda_1 = \lambda_2 = \dots = \lambda_{n+1} = \lambda < (n+1)^{-1/n}$, то $\mu(A) = 0$. В частности, для СИФ S на плоскости ($n = 2$) аттрактор A имеет нулевую меру Лебега при $\lambda < 1/\sqrt{3}$.

В [1] авторами доказана следующая теорема.

Теорема 2. Пусть $S = \{f_i | i = 1..n+1\}$ – система итерированных функций в евклидовом пространстве R^n , состоящая из $n+1$ гомотетии с коэффициентами подобия $\lambda_i \in (0,1)$ и аффинно независимыми центрами A_1, A_2, \dots, A_{n+1} . Если

$$\lambda_1 + \lambda_2 + \dots + \lambda_{n+1} \geq n,$$

то аттрактором A является симплекс с вершинами в точках A_1, A_2, \dots, A_{n+1} .

Из теоремы 2 вытекает сразу следующее утверждение.

Следствие 2. Если $\lambda_1 + \lambda_2 + \dots + \lambda_{n+1} \geq n$, мера Лебега аттрактора A СИФ S положительна и равна n -мерному объему симплекса с вершинами в точках A_1, A_2, \dots, A_{n+1} .

В случае равенства коэффициентов подобия получаем результат *Бромхеда, Монталди и Сидорова [1], упомянутый выше*.

Следствие 3. Если $\lambda_1 = \lambda_2 = \dots = \lambda_{n+1} = \lambda \geq n/(n+1)$, то мера Лебега аттрактора A СИФ S положительна и равна n -мерному объему симплекса с вершинами в центрах гомотетий A_1, A_2, \dots, A_{n+1} . В частности, для СИФ S на плоскости ($n = 2$) при $\lambda \geq 2/3$ мера Лебега аттрактора A положительна и равна площади треугольника с вершинами в центрах гомотетий.

В [1] авторами доказано следующее утверждение.

Теорема 3. Пусть $S = \{f_i | i = 1..n+1\}$ – система итерированных функций в евклидовом пространстве R^n , состоящая из $n+1$ гомотетии с коэффициентами подобия $\lambda_i \in (0,1)$ и аффинно независимыми центрами A_1, A_2, \dots, A_{n+1} . Если

$$\lambda_i + \lambda_j \leq 1, \quad \forall i \neq j, \quad i, j = 1 \dots n+1, \quad (2)$$

то аттрактором A является канторово множество в симплексе с вершинами в точках A_1, A_2, \dots, A_{n+1} , причем размерности Минковского и Хаусдорфа аттрактора A равны d , где d – единственное решение уравнения $\lambda_1^d + \lambda_2^d + \dots + \lambda_{n+1}^d = 1$.

Следствие 4. Если выполнены условия (2) теоремы 3, то n -мерная мера Лебега аттрактора A СИФ S равна нулю.

Доказательство. Обозначим через G компактное выпуклое множество, заданное неравенствами $\lambda_i + \lambda_j \leq 1, \forall i \neq j, \lambda_i \geq 0, i, j = 1 \dots n + 1$. Рассмотрим на G функцию

$$d(\lambda_1, \lambda_2, \dots, \lambda_{n+1}) = \sqrt{\lambda_1^2 + \lambda_2^2 + \dots + \lambda_{n+1}^2},$$

равную расстоянию от начала координат до точки $(\lambda_1, \lambda_2, \dots, \lambda_{n+1}) \in G$. Нетрудно видеть, что наибольшее значение этой функции, равное 1, достигается только в точках M_i с координатами $\lambda_j = 0, \forall j \neq i, \lambda_i = 1, i, j = 1 \dots n + 1$. Следовательно,

$$\lambda_1^2 + \lambda_2^2 + \dots + \lambda_{n+1}^2 \leq 1$$

в G . Поскольку для коэффициентов гомотетии $0 < \lambda_i < 1, i = 1 \dots n + 1$, то

$$\lambda_1^2 + \lambda_2^2 + \dots + \lambda_{n+1}^2 < 1.$$

Так как функция $f(s) = \lambda_1^s + \lambda_2^s + \dots + \lambda_{n+1}^s$ является убывающей, то

$$\lambda_1^n + \lambda_2^n + \dots + \lambda_{n+1}^n \leq \lambda_1^2 + \lambda_2^2 + \dots + \lambda_{n+1}^2 < 1$$

для $n \geq 2$. Таким образом, выполнено условие (1) теоремы 1 и, следовательно, n -мерная мера Лебега аттрактора A СИФ S равна нулю.

Пример

Пусть $\lambda_1 = 0.5, \lambda_2 = 0.6, \lambda_3 = 0.4$. Поскольку $\lambda_1^2 + \lambda_2^2 + \lambda_3^2 = 0.77 < 1$, то согласно теореме 1 мера Лебега аттрактора A СИФ $S = \{f_1, f_2, f_3\}$ равна нулю. Визуализация аттрактора A СИФ S реализована с помощью рандомизированного алгоритма в программе Махіма.

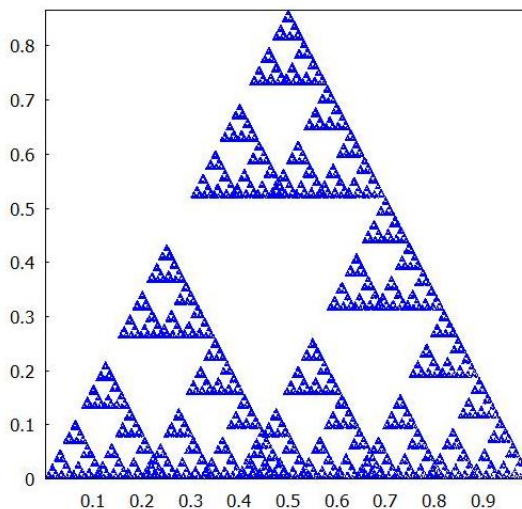


Рис. 1. Аттрактор A СИФ S при $\lambda_1 = 0.5, \lambda_2 = 0.6, \lambda_3 = 0.4$.

Заключение

Для системы итерированных функций $S = \{f_i | i = 1..n + 1\}$, состоящей из гомотетий в n -мерном евклидовом пространстве R^n с аффинно независимыми центрами A_i и произвольными коэффициентами подобия $\lambda_i \in (0,1), i = 1..n + 1$, найдены достаточные условия, при которых мера Лебега аттрактора A СИФ S : 1) положительна; 2) равна нулю.

Библиографический список

1. Broomhead D. Golden gaskets: variations on the Sierpiński sieve / D. Broomhead, J. Montaldi, N. Sidorov // *Nonlinearity*. – 2004. – V. 17. – P. 1455–1480.
2. Simon K. On the dimension of self-similar sets / K. Simon, B. Solomyak // *Fractals*. – 2003. – V. 10. – P. 59–65.
3. Jordan Th. Dimension of fat Sierpiński gaskets / Th. Jordan // *Real Anal. Exchange*. – 2005. – V. 31. – P. 97–110.
4. Багаев А.В. Аттракторы систем трех итерированных гомотетий евклидовой плоскости // А.В. Багаев, А.В. Киселева / Сборник материалов XXIX Всероссийской научно-практической конференции по графическим информационным технологиям и системам «КОГРАФ – 2019», 18 апреля 2019 года, Н. Новгород, 2019. – С.136-140.
5. Багаев А.В. О многомерных аналогах треугольника Серпинского // А.В. Багаев, А.В. Киселева / Сборник материалов XXVI Международной научно-технической конференции "Информационные системы и технологии - 2020", 24, 27-28 апреля 2020 года, Н. Новгород, 2020. - С. 1148-1152.

A.V. Bagaev, A.V. Kiseleva

ON THE LEBESGUE MEASURE OF THE ATTRACTORS GIVEN BY HOMOTETIES WITH AFFINE INDEPENDENT CENTERS

Nizhny Novgorod state technical university n.a. R.E. Alekseev

We consider the attractors of iterated function systems in n -dimensional Euclidean space consisting of $n + 1$ homotheties with affine independent centers and arbitrary similarity coefficients $\lambda_i \in (0,1)$. We found the conditions under which the Lebesgue measure of such attractors is zero or positive.

Keywords: iterated function system, attractor, fractal, Lebesgue measure, Cantor set

О РАЗРАБОТКЕ ВЕБ-ПРИЛОЖЕНИЯ ДЛЯ ВЗАИМОДЕЙСТВИЯ С БАЗОЙ ДАННЫХ ЛАБОРАТОРНЫХ ЭКСПЕРИМЕНТОВ В ГИДРОДИНАМИЧЕСКОМ ЛОТКЕ

Нижегородский государственный технический университет им. Р.Е. Алексева

В работе представлено веб-приложение, которое предназначено для взаимодействия с базой данных, содержащей информацию о лабораторных экспериментах по генерации и распространению внутренних и поверхностных волн в гидродинамическом лотке научно-исследовательской лаборатории Моделирования природных и техногенных катастроф НГТУ им. Р.Е. Алексева.

Ключевые слова: веб-приложение, база данных, графический пользовательский интерфейс, лабораторные эксперименты.

Введение

В научно-исследовательской лаборатории Моделирования природных и техногенных катастроф (НИЛ МПиТК) проводятся эксперименты по генерации и распространению внутренних и поверхностных волн в гидродинамическом лотке. На основе свободной реляционной системы управления базами данных MySQLами была создана база данных (БД), содержащая информацию о них [1,2]. Взаимодействие с базой до недавнего времени было организовано с помощью инструмента для визуального проектирования БДМ у SQLWorkbench. К его недостаткам можно отнести неудобную для рядового пользователя, не знакомого с SQL, систему запросов, а также невозможность удаленной работы с данными лабораторных экспериментов. В связи с этим было принято решение о разработке веб-приложения, не требующего от конечного пользователя каких-либо навыков программирования и позволяющего взаимодействовать с БД дистанционно. Описание базы данных лабораторных экспериментов в гидродинамическом лотке и её EER-диаграммы подробно рассмотрено в [3].

Описание веб-приложения

Веб-приложение было разработано на основе бесплатного фреймворка с открытым исходным кодом Django, написанного на языке Python.

Python – интерпретируемый высокоуровневый язык программирования. Он прост в освоении, так как, по сравнению с другими высокоуровневыми языками, максимально приближен к синтаксису английского языка. Основными преимуществами Python являются динамическая типизация и автоматическое выделение памяти. Python – очень популярный язык программирования как среди непрофессиональных разработчиков, так и среди крупных корпораций таких, как Google, Facebook и Django Software Foundation. Эти компании использовали его для создания таких проектов, как YouTube, Instagram и упомянутый выше Django.

Django – свободный фреймворк для разработки веб-приложений. Его главными особенностями являются универсальность, безопасность, масштабируемость. Универсальность заключается в том, что при помощи данного модуля можно реализовать практически любой веб-сайт – от социальных сетей до новостных лент. Он может работать с любой клиентской структурой и обрабатывать данные в большинстве форматов, таких как JSON, XML и прочие. Django обладает встроенной защитой против большинства типов CSRF атак, внедрения SQLи от «скликивания». Архитектура фреймворка имеет модульную составляющую, что позволяет свободно убирать ненужные компоненты при разработке и добавлять новые [4].

Приложения, разработанные на базе фреймворка Django, являются кроссплатформенными, то есть они способны работать с несколькими аппаратными платформами или операционными системами.

Также к фреймворку Django можно легко подключить панель администратора, которая и позволила нам быстро и достаточно просто создавать графический интерфейс для взаимодействия пользователя с базой данных лабораторных экспериментов. В разработанном нами веб-приложении можно просматривать записи, содержащиеся в любой из таблиц БД, однако мы остановимся только на одной из них, Experiment, так как эта таблица имеет доступ ко всем остальным (это происходит за счет «внешних» ключей базы данных, то есть в этой таблице содержится идентификационный номер записи из другой таблицы).

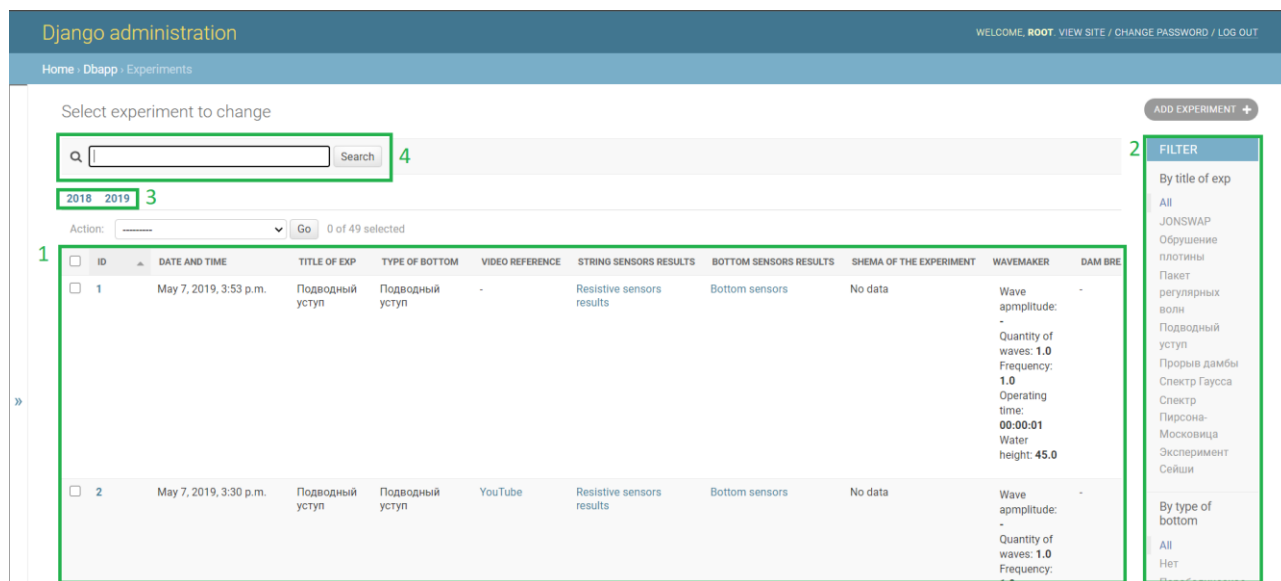


Рис. 1. Пример графического интерфейса веб-приложения (1 – таблица данных, 2 – таблица фильтров, 3 – список фильтров по дате проведения экспериментов, 4 – окно поиска)

Графический интерфейс (рис. 1) состоит из следующих фрагментов: таблица данных, таблица фильтров, список фильтров по дате проведения экспериментов, окно поиска.

Таблица данных отображает информацию о проведенных экспериментах. Если пользователя не устраивает порядок сортировки, то его можно изменить, нажав на заголовок соответствующей колонки.

Таблица фильтров позволяет отсортировать данные по интересующим пользователя параметрам.

Окно поиска предоставляет возможность производить поиск по тем параметрам, которые не указаны в списке фильтров, например, по плотности жидкости, используемой в эксперименте.

На рис. 2 представлен пример поиска записей по названию эксперимента.

2018 December 3 December 4

Action: [dropdown] Go 0 of 2 selected

ID	DATE AND TIME	TITLE OF EXP	TYPE OF BOTTOM	VIDEO REFERENCE	STRING SENSORS RESULTS	BOTTOM SENSORS RESULTS	HEMA OF THE EXPERIMENT	WAVEMAKER	DAM BREAK EX
15	Dec. 3, 2018, 3:46 p.m.	JONSWAP	Нет	YouTube	Resistive sensors results	Bottom sensors	No data	Wave amplitude: 5.0 Quantity of waves: - Frequency: 0.95 Operating time: - Water height: 40.0	-
16	Dec. 4, 2018, 11:05 a.m.	JONSWAP	Нет	YouTube	Resistive sensors results	Bottom sensors	No data	Wave amplitude: 10.0 Quantity of waves: - Frequency: 0.95 Operating time: - Water height: 40.0	-

By title of exp

All

JONSWAP

Обрушение плотины

Пакет регулярных волн

Подводный уступ

Прорыв дамбы

Спектр Гаусса

Спектр

Пирсона-Московица

Эксперимент

Сейши

By type of bottom

All

Нет

Параболическое дно

Подводный

Рис. 2. Пример поиска записи по названию эксперимента

Чтобы открыть меню редактирования записи, необходимо нажать на соответствующий эксперименту идентификатор. Окно редактирования представлено на рис.3.

Django administration

Home > Dbapp > Experiments > Experiment object (1)

Change experiment

Date and time: Date: 2019-05-07 Today | 📅
Time: 15:53:00 Now | 🕒
Note: You are 3 hours ahead of server time.

Title of exp:

Type of bottom:

Video reference:

Result: 🗑️ +

Schema of exp reference:

Type of forming: 🗑️ +

Type of wave:

Duration of the exp: Now | 🕒
Note: You are 3 hours ahead of server time.

Type of stratification: 🗑️ +

Polarity:

Laser: 🗑️ +

Types of experiment: 🗑️ +

Рис. 3. Окно редактирования записи

Жирным шрифтом обозначены параметры, которые подсоединяются к таблице Experiment по внешнему ключу. Для их редактирования необходимо нажать на «карандаш», а для создания новой записи – на знак «плюс». При этом откроется новое окно, позволяющее работать с записью в той таблице, к которой относится этот внешний ключ (рис. 4).

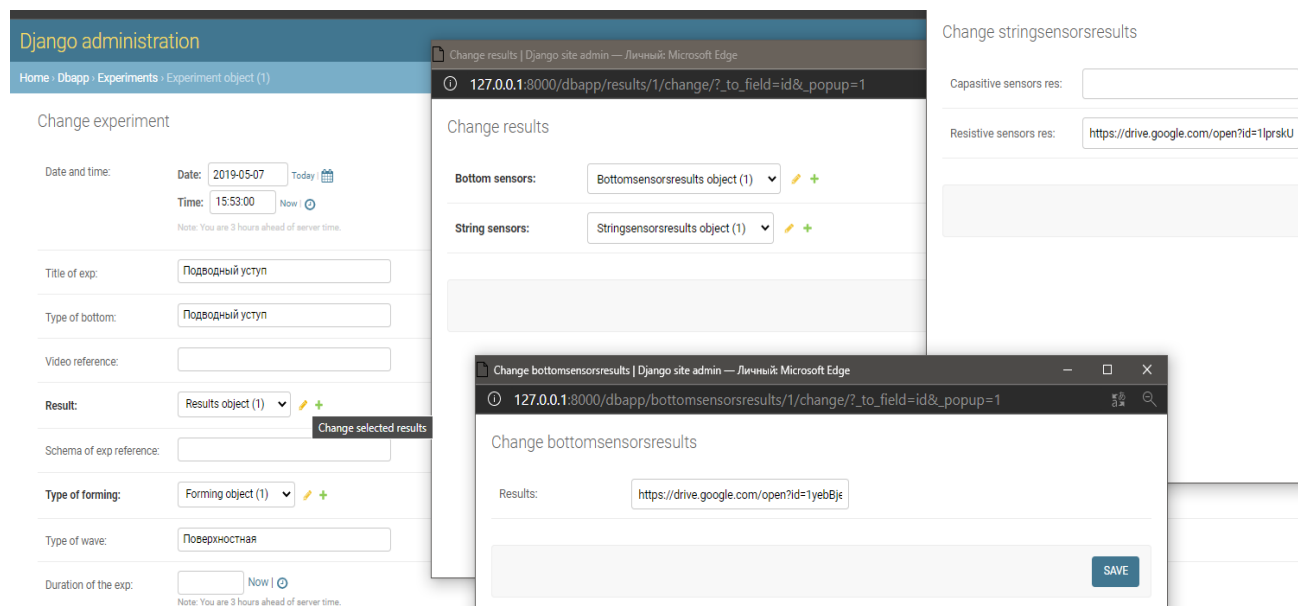


Рис. 4. Окно редактирования записи по внешнему ключу

Выводы

В настоящей работе представлено веб-приложение для работы с базой данных лабораторных экспериментов по генерации и распространению внутренних и поверхностных волн в гидродинамическом лотке НИЛ МПиТК НГТУ им. Р.Е. Алексева, разработанное на основе фреймворка Django. Приведено описание интерфейса приложения и его основные возможности: просмотр, редактирование добавление новых записей, а также поиск данных по заданным параметрам.

Представленные результаты получены при поддержке гранта Президента РФ по государственной поддержке ведущих научных школ РФ НШ-2485.2020.5.

Библиографический список

1. Куркин А.А. База данных лабораторных экспериментов по генерации и распространению внутренних волн / А.А. Куркин, О.Е. Куркина, Н.А. Санников, Л.В. Талалушкина, А.В. Рыбин, А.С. Епифанова // Свидетельство о государственной регистрации базы данных № 2020621291 от 24 июля 2020 г.
2. Куркин А.А. База данных лабораторных экспериментов по генерации и распространению поверхностных волн / А.А. Куркин, О.Е. Куркина, Н.А. Санников, Н.М. Лиходеев, Л.В. Талалушкина, А.В. Рыбин, А.С. Епифанова // Свидетельство о государственной регистрации базы данных № 2020621342 от 03 августа 2020 г.
3. Епифанова А.С. База данных лабораторных экспериментов по генерации и распространению внутренних и поверхностных волн. Первые результаты / А.С. Епифанова, О.Е. Куркина, Н.М. Лиходеев, А.Р. Гиниятуллин, Н.А. Санников, Л.В. Талалушкина // Сборник материалов 30-й Всероссийской научно-практической конференции по графическим информационным технологиям и системам. – 2020. – С. 302-306.
4. Дронов В.А. Django 2.1. Практика создания веб-сайтов на Python / В.А. Дронов. – СПб.: БХВ-Петербург, 2019. – 672 с.

N.A. Sannikov, L.V. Talalushkina, A.S. Epifanova

ABOUT THE DEVELOPMENT OF A WEB-APPLICATION FOR INTERACTING WITH DATABASE OF LABORATORY EXPERIMENTS IN A HYDRODYNAMIC TRAY.

Nizhny Novgorod state technical university n. a. R.E. Alekseev

Purpose: This article describes the web-application for working with the database of laboratory experiments on the generation and propagation of internal and surface waves in the hydrodynamic wavetank.

Design/methodology/approach: We use the free open-source framework Django written in the high-level programming language Python. Our application interacts with a database created on the basis of the free relational database management system MySQL.

Findings: The developed web application allows the user to view, add, correct and delete data from the database of laboratory experiments.

Originality/value: The web-application help researchers in the field of studying the dynamics of surface and internal waves.

Keywords: web application, database, graphical user interface, laboratory experiments.

РАЗРАБОТКА И СОЗДАНИЕ ИНФОРМАЦИОННОГО ПОРТАЛА ДИССЕРТАЦИОННОГО СОВЕТА С ИСПОЛЬЗОВАНИЕМ ФРЕЙМВОРКА DJANGO

Нижегородский государственный технический университет Р.Е. Алексеева

В статье представлена концепция информационного портала диссертационного совета, а также разработана структурная схема ресурса и определено содержание его разделов. На основании анализа современных технологий веб программирования и требований к проектируемому сайту сделан выбор в пользу языка python и фреймворка Django. Определена структура баз данных. Реализована первая версия информационного ресурса.

Ключевые слова: веб-приложение, фреймворк, python, база данных

В современном мире сложно переоценить влияние сети интернет на все сферы человеческой жизни – от маркетинга и продвижения товаров и услуг до науки и образования. В духе времени создание «электронных» представительств как производственных предприятий, так и научных, научно-образовательных институтов и их подразделений. Веб-сайт – это удобное и доступное средство информирования и взаимодействия с целевой аудиторией, позволяющее решать ряд важных задач, таких как популяризация актуальных научных исследований и проектов, а также взаимодействия между различными исследовательскими группами; в то же время выход организации на «информационную» площадку позволяет нивелировать пространственные и временные ограничения деятельности.

Наряду с кафедрами и научными лабораториями в некоторых высших учебных заведениях и научных институтах существуют также диссертационные советы, которые функционируют под руководством Высшей аттестационной комиссии, и, несомненно, повышают статус организации, на базе которой они созданы, поскольку в их состав входят ведущие специалисты по представленной специальности (специальностям). Деятельность совета по защите кандидатских/докторских диссертаций строго регламентирована нормативно-правовыми актами (см., например, [1, 2]) и сопряжена с оформлением существенного количества документов по предстоящим и проведенным защитам диссертаций, а также подготовкой отчетов (по деятельности совета за определенный период). Представляется важным не только решить задачу максимальной автоматизации некоторых аспектов деятельности совета, но также и способствовать предоставлению актуальной и достоверной информации о «жизни» совета соискателям ученых степеней и заинтересованному научному сообществу.

Целью настоящей работы является создание информационного портала для диссертационного совета Д212.165.10, созданного на базе Нижегородского государственного технического университета им. Р.Е. Алексеева.

В настоящее время существует множество средств и инструментов для создания специализированного веб-сайта. В зависимости от предпочитаемого языка программирования можно также использовать тот или иной фреймворк. Фреймворк (от англ. framework — каркас) — это программная библиотека, реализующая большую часть типовой функциональности разрабатываемого продукта [3] (Django и Flask для Python, ASP.Net для C#, Angular для JS и др.).

Важным этапом проектирования веб-ресурса является разработка его функциональной схемы. На рис. 1 представлена структурная схема информационного портала диссертационного совета.

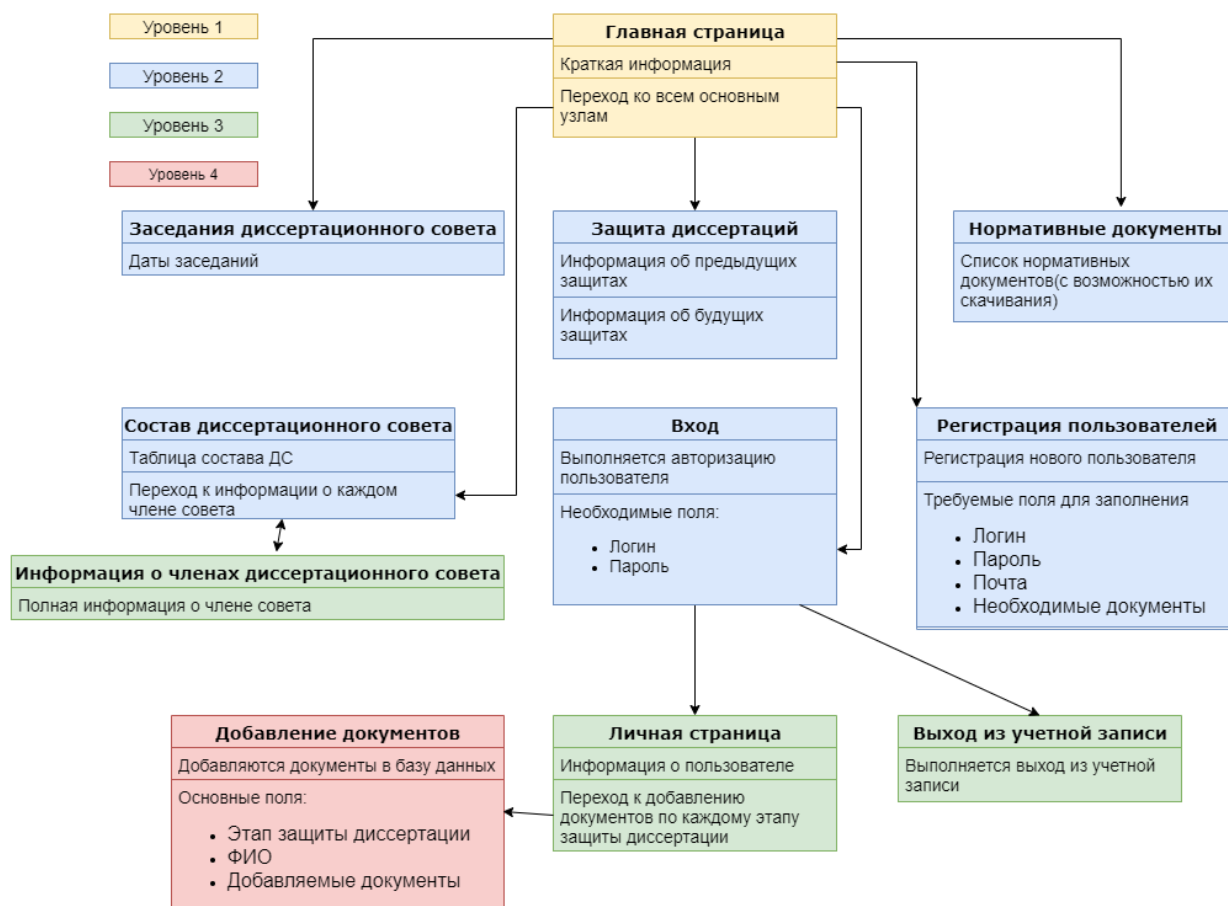


Рис. 1. Функциональная схема информационного портала диссертационного совета.

Как видно из рис.1, разработанная концепция сайта имеет традиционную трехуровневую структуру, включающую пользовательский интерфейс, средства доступа к информации и средства обработки запросов.

В приложении для диссертационного совета были реализованы следующие функции:

- регистрация и авторизация пользователей;
- доступ к базе данных нормативных документов;
- доступ к базе данных, содержащей сведения о публикационной активности, результатах интеллектуальной собственности и участии в конференциях и семинарах для всех членов диссертационного совета;
- загрузка и отображение документов по каждому этапу дела о защите кандидатской/докторской диссертации;
- настройка административной панели;
- настройка редактирования и добавления новостей и информационных сообщений;
- разделение пользователей на группы с различными правами доступа.

Для организации хранения и доступа к информации реализованы базы данных, схемы которых представлены на рис.2.

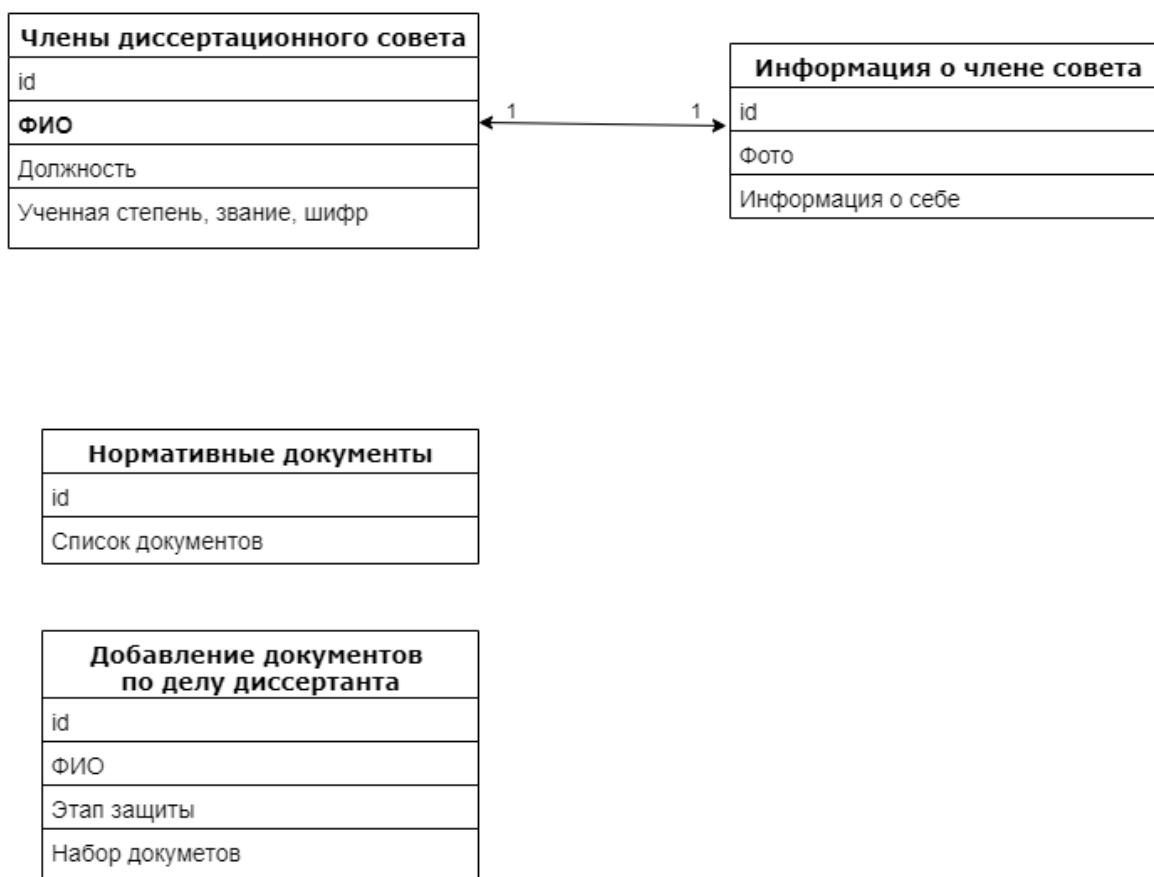


Рис. 2. Схема базы данных и таблиц.

Для программной реализации системы выбран Python – современный высокоуровневый язык программирования, имеющий свободное распространение. Он обладает такими возможностями, как динамическая типизация, поддержка ООП (объектно-ориентированное программирование), а также кроссплатформенность. При этом он имеет минималистический синтаксис, что позволяет легко писать код, а в будущем и дополнять его [5]. Этот язык подходит для решения множества задач, в том числе и написания веб-приложения. Совместно с Python могут использоваться два самых популярных фреймворка – Django и Flask. Мы сделали выбор в пользу Django, так как API-интерфейсы данного фреймворка в отличие от Flask предоставляют возможности для управления всеми частями проекта, включая базу данных и элементы шаблона. Использование уже готовой административной панели упрощает создание приложения [4].

Django – это высокоуровневый веб-фреймворк, основные преимущества которого заключаются в следующих функциональных возможностях:

- использовании архитектуры MVC (model-view-controller), которая отвечает за разделение интерфейса и логики;
- наличии удобного отладочного веб-сервера, развитой миграции, а также готовой административной панели;
- легкости в добавлении дополнительных библиотек;

выполнении таких задач, как связь с базой данных (сохранение выгруженных файлов выполняется самим фреймворком) [3, 8].

В фреймворке Django для работы с базой данных использовался реляционный sql-язык SQLite. Он является встраиваемой СУБД, который обращается к файлам хранения напрямую. Данный подход помогает уменьшить время отклика, и упрощает программуSQLiteхранит всю базу данных в одном файле на компьютере, на котором выполняется программа. Так же данная СУБД позволяет нескольким процессам одновременно читать данные из одной базы. Запись может выполняться только если в данный момент нет никаких других запросов [6, 7].

В Django взаимодействие с базой данных происходит по ORM (Object-RelationalMapping) – технологии программирования, предназначенной для связи базы данных с концепциями объектно-ориентированных языков программирования. Результатом программной реализации проекта стала первая версия информационного портала диссертационного совета, интерфейс которой представлен на рис. 3.

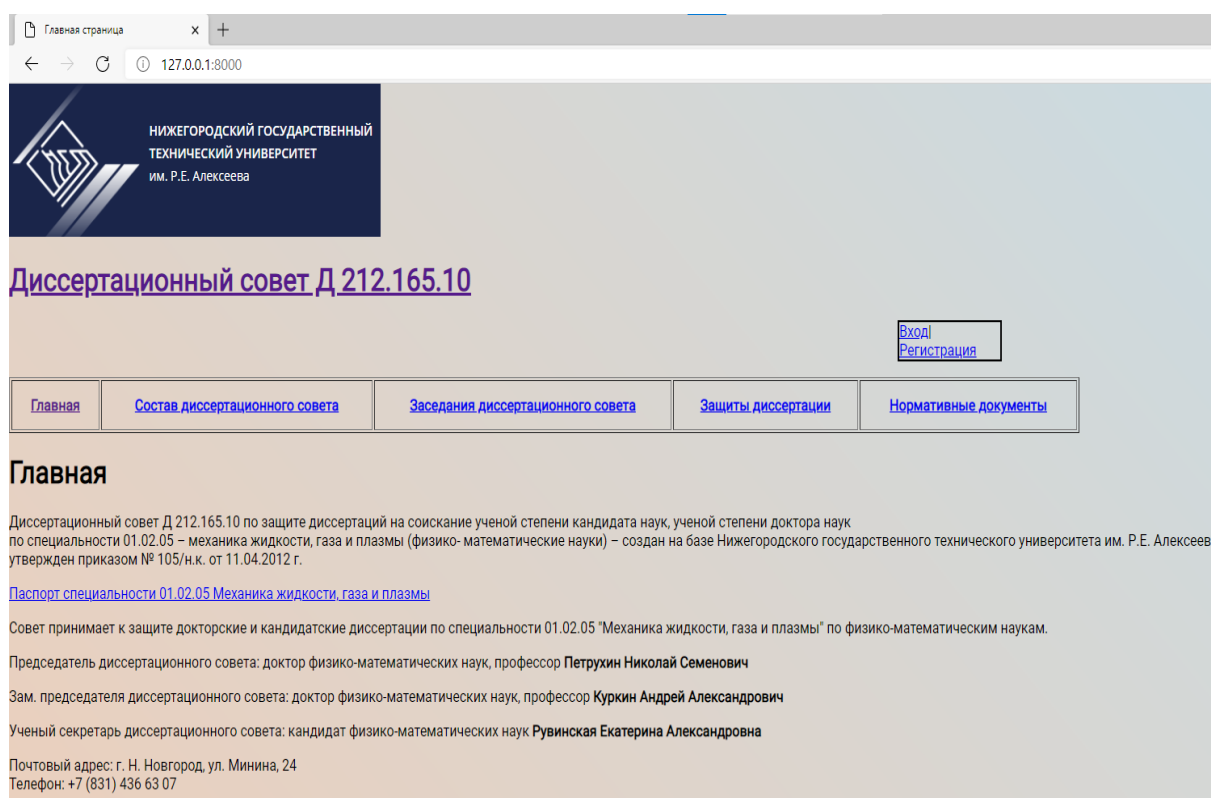


Рис. 3. Интерфейс сайта диссертационного совета

Таким образом, при создании информационного портала диссертационного совета были изучены современные подходы и технологии разработки, проанализированы их преимущества и недостатки. Спроектирована и реализована первая версия веб-приложения с возможностью дальнейшей доработки функционала. Планируется обновление дизайна интерфейса портала, а также добавление блока загрузки информации о публикациях членов совета за выбранных период из доступных электронных библиотек.

Представленные результаты получены при поддержке гранта Президента РФ по государственной поддержке ведущих научных школ РФ НШ-2485.2020.5 и стипендии Президента Российской Федерации молодым ученым и аспирантам (СП-1225.2019.5).

Библиографический список

1. Постановление Правительства Российской Федерации от 24 сентября 2013 г. N 842 г. Москва "О порядке присуждения ученых степеней"
2. Приказ Министерства образования и науки РФ от 10 ноября 2017 г. № 1093 "Об утверждении Положения о совете по защите диссертаций на соискание ученой степени кандидата наук, на соискание ученой степени доктора наук"
3. Владимир Дронов. Django 3.0. Практика создания Web-сайтов на Python. - СПб.: БХВ-Петербург, 2021. - 704с.
4. David Ashley. Foundation Dynamic Web Pages with Python. Austin, TX, USA, 2020.
5. Ахметшин Д.А. Разработка приложений на языке программирования Python с использованием Фреймворка Django. - Казань: редакционно-издательский центр «Школа», 2019. – 116 с.
6. Allen G., Owens M. SQL for SQLite. The Definitive Guide to SQLite, pp. 47-86.
7. Березуцкая Л.А. SQLite main commands and their comparison to SQL for an android app. // Сборник материалов Всероссийской молодежной научно-практической конференции с международным участием «LinguaNet». Под общ. ред. Ю.А. Иванцовой, Н.С. Руденко. 2019
8. Рябова К.М. Фреймворк Django: Архитектура и возможности. //Современные технологии: актуальные вопросы, достижения и инновации, сборник статей XIII Международной научно-практической конференции: в 2 частях. 2018

I.I. Semerenko, E.A. Rouvinskaya

DEVELOPMENT AND CREATION OF THE INFORMATION PORTAL OF THE DISSERTATION COUNCIL USING THE DJANGO FRAMEWORK

Nizhny Novgorod state technical university n. a. R. E. Alekseev

The article presents the concept of the information portal of the dissertation council, as well as a structural diagram of the resource. The content of portal's sections is determined. We use the python language and the Django framework. Our choice is based on the analysis of modern web programming technologies and requirements for the designed site. The structure of the databases is defined. The first version of the information resource has been implemented.

Keywords: web application, framework, python, database

ОПТИМИЗАЦИЯ ПРОЦЕССА ОБРАБОТКИ ИНФОРМАЦИИ В ДЕТСКОЙ ПОЛИКЛИНИКЕ

*Нижегородский государственный технический университет им. Р.Е. Алексеева
ФГБОУ ВО "ПИМУ" Минздрава России*

В работе описан программный комплекс, обеспечивающий оптимизацию процесса обработки информации в условиях детской поликлиники. Разработанное программное обеспечение нацелено на увеличение эффективности обработки полученной информации и создания базы данных для проведения анализа.

Ключевые слова: база данных, программный комплекс, детская поликлиника, информация.

Введение

В настоящее время информационная индустрия связана с созданием технических методов, средств и технологий для получения и хранения имеющейся информации. Учреждения здравоохранения различного типа не являются исключением, нуждаясь в создании удобных для пользователей и надёжных баз данных. Управление данными для каждого врача может быть использовано для получения информации о пациенте, обеспечит её пополнение и проведения сравнительного анализа.

Целью выполнения работы является создание программного комплекса для оптимизации обработки информации в условиях детской поликлиники. Разработанная программа должна представлять собой программное средство хранения и обработки информации, а также получения отчётов в требуемом виде. Созданная информационная система способна повысить эффективность деятельности врача на основе автоматизации трудоемких операций обработки статистического анализа информации.

Создаваемый программный комплекс обеспечивает реализацию следующих функций: авторизация пользователей; хранение данных; обработка и редактирование данных; выгрузка информации о хранящихся данных в требуемом формате; удалённый доступ к данным.

Разработка программного комплекса

Сравнительный анализ популярных языков программирования позволил выбрать для программной реализации язык python как наиболее удобное и компактное средство для создания приложений, работающих с большими объемами данными. На сегодняшний день python является самым распространенным языком для построения научных и технических программных продуктов ввиду компактности и наглядности синтаксиса языка, мощной технической поддержки и наличия крупнейшей свободной базы библиотек (пакетов) с реализованными алгоритмами практически во всех областях научной деятельности. Язык программирования python поддерживает структурное, объектно-ориентированное, функциональное, императивное и аспектно-ориентированное программирование. Основные его архитектурные черты – это динамическая типизация, автоматическое управление памятью, механизм обработки исключений, поддержка многопоточных вычислений, высокоуровневые структуры данных. При разработке программного кода мы использовали объектно-ориентированный подход, который в полной мере поддерживается языком программирования python. К плюсам языка python также относятся его простота, большое количество библиотек для использования в самых различных областях, бесплатное распространение и широкая поддержка сообществом программистов.

Графический интерфейс программы будет разрабатываться с помощью библиотеки `pyQt`, являющейся оберткой к популярной C++ графической библиотеке `Qt`. Данная библиотека предоставляет удобную среду для создания оконных приложений с широким набором готовых компонентов пользовательского интерфейса (виджетов), в то же время допускающих гибкую настройку под непосредственные нужды разработчика. Непосредственная математическая обработка данных будет производиться посредством популярного специализированного `python` модуля для научных вычислений `numpy`. Данный модуль позволяет проводить с высокой эффективностью математические операции над многомерными массивами, включая частные случаи одномерных массивов (векторы) и двумерных массивов (матрицы).

Отображение данных в форме графиков и диаграмм будет осуществляться с использованием пакета `matplotlib`. Данный модуль является наиболее распространенным средством для построения графиков в языке `python` с возможностью чрезвычайно гибкой настройки и отображения различных типов графиков и диаграмм (график кривой, гистограмма, столбчатая и круговая диаграмма, поверхность в трехмерном пространстве).

Наконец, для удобной организации программного кода и обеспечения возможности сборки проекта в исполняемый `exe` файл мы будем использовать специализированную `python` библиотеку `fb`, непосредственно предназначенную для создания `pyQt` проектов и их сборки в исполняемый файл для любых операционных систем.

В последние годы при разработке приложений (в особенности `web` приложений) приобрел популярность так называемый `serverless` подход, при котором вместо самостоятельной настройки собственного физического или облачного сервера используется уже сконфигурированная стандартная система, позволяющая производить автоматическое развертывание.

Одним из наиболее распространенных `serverless` решений является сервис `firebase` от `google`, который включает хостинг, систему аутентификации пользователей посредством адреса электронной почты и/или мобильного телефона, облачное хранилище файлов `firestorage`, базу данных реального времени `firestore` и множество других возможностей. В данной работе будет использоваться удобная `noSQL` база данных `firestore`.

Для хранения данных будет использоваться `nosql` база данных реального времени `firestore`. Данная база данных предполагает организацию информации в иерархической форме, состоящей из документов, объединенных в коллекции, при этом каждый документ может содержать как непосредственно данные, так и вложенные коллекции документов. Облачный сервис `firebase` планируется использовать для хранения данных программного приложения, при этом для доступа к сервису из `python` коды мы воспользуемся стандартным `firebase`-менеджером – `python` пакет `firebase-admin`.

Графический пользовательский интерфейс созданной медицинской информационной системы (`hospimedbase`) в состоянии после запуска программы, ввода логина и пароля пользователя и загрузки тестовой базы данных представлен на рисунке 1. На данном рисунке присутствуют следующие его основные составные части: программное меню, содержащее возможные пользовательские операции (добавление, удаление и редактирование информации о пациентах, включая добавление медосмотров); компактное программное меню, составленное из стилизованных кнопок, дублирующее программное меню; таблица, содержащая список пациентов, сохраненных в базе данных и доступных для текущего пользователя; форма, содержащая фильтры по пациентам в базе данных; информационная панель программы, отображающая текстовые сообщения о результатах основных программных действий.

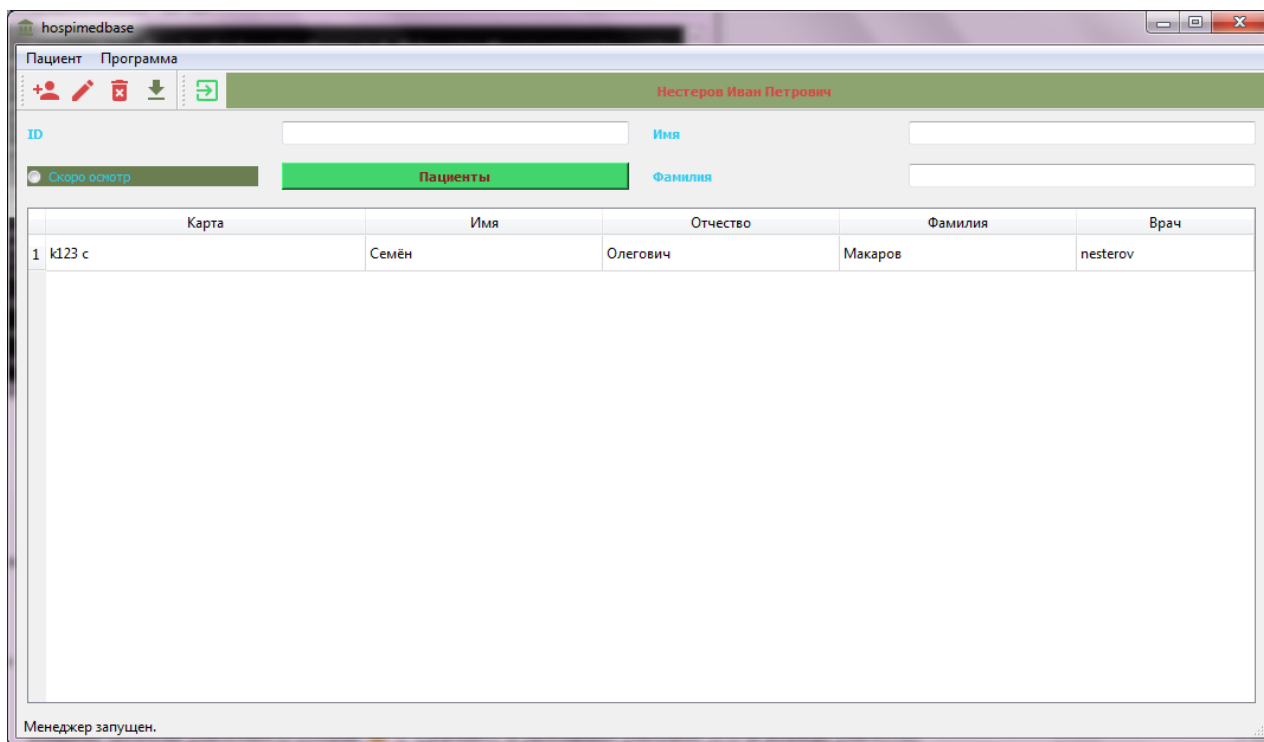


Рис. 1. Пользовательский интерфейс программы

На рисунке 2 представлена форма для авторизации в программе. После ввода логина и пароля пользователя и нажатия на кнопку подтверждения осуществляется запрос к базе данных, и в случае корректности введенной информации осуществляется вход в программу с интерфейсом пользователя, представленным на рисунке 3 и интерфейсом администратора. Отметим, что при нажатии на кнопку «Пациенты» на рисунке 3, происходит смена режима отображения, и в таблице начинает отображаться список пользователей системы, как показано на рисунке 4.

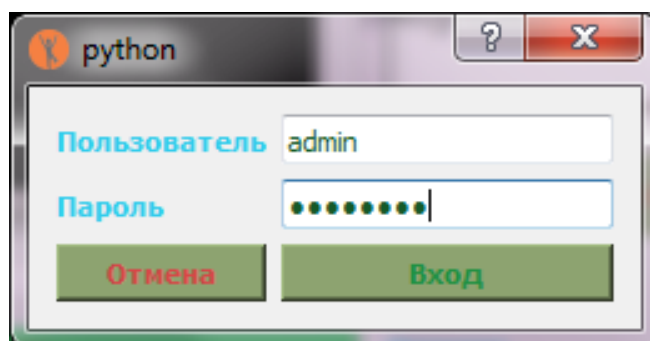


Рис. 2. Форма для авторизации в программе.

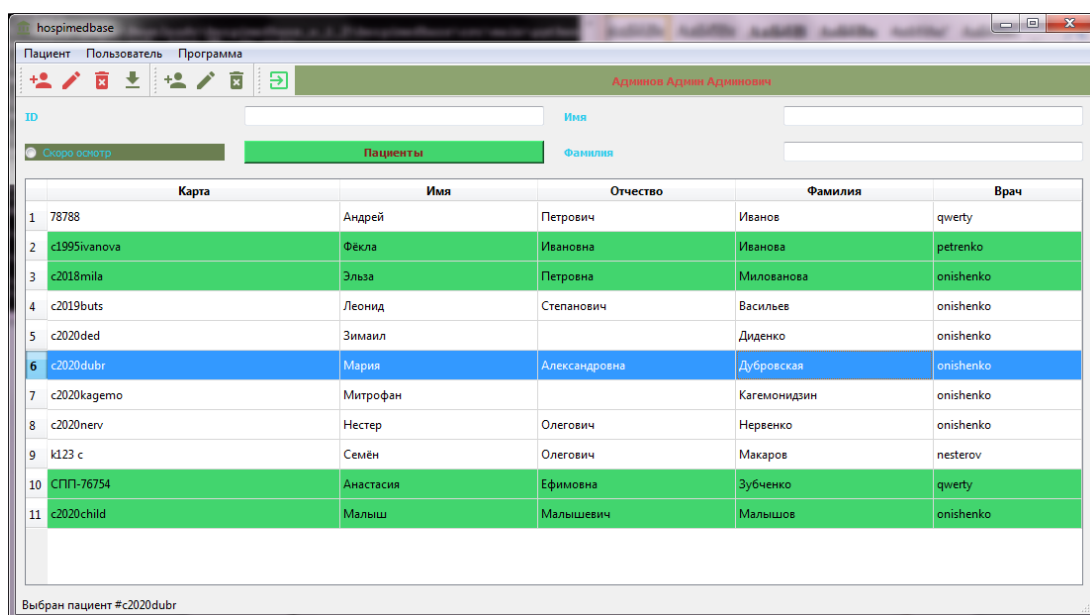


Рис. 3. Пользовательский интерфейс программы (администратор). Отображается полный список пациентов, также доступны операции с пользователями (врачами)

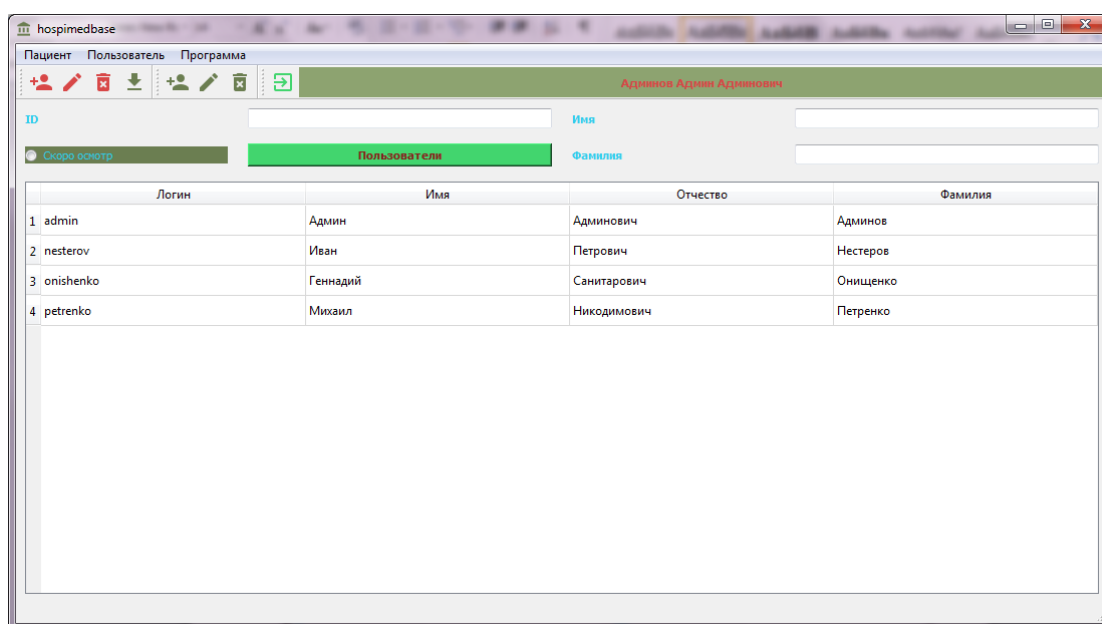


Рис. 4. Отображение текущего списка пользователей для администратора программы

Выводы

Таким образом, разработанный программный комплекс способен повысить эффективность деятельности врача детской поликлиники на основе автоматизации трудоемких операций хранения большого количества информации о пациентах, её обработки, выборки и вывода информации в необходимом формате, проведения статистической обработки и дальнейшего анализа полученных данных. В дальнейшем планируется усовершенствование разрабатываемого программного обеспечения в виде улучшения системы уведомлений путём создания отдельной коллекции в базе данных и добавление в программу возможности отслеживания изменений в содержании базы данных в режиме реального времени.

Библиографический список

1. Доусон М. Программируем на Python. – СПб.: Питер, 2014. – 416 с.
2. Лутц М. Программирование на Python, том II, 4-е издание. – Пер. с англ. – СПб.: Символ-Плюс, 2011. – 992 с.
3. Шапошникова С. Основы программирования на Python. Вводный курс.
4. Чаплыгин А.Н. Учимся программировать вместе с питоном.
5. Шварц Б., Зайцев П., Ткаченко В. и др. MySQL. Оптимизация производительности (2-е издание)
6. Посыпкина Е.В., Чекалова С.А. [Программный комплекс для комплексной оценки состояния здоровья детей в амбулаторно-поликлинической службе.](#) Сборник материалов XXVI Международной научно-технической конференции. Нижегородский государственный технический университет им. Р.Е. Алексева. 2020. С. 1194-1197.
7. Посыпкина Е.В., Чекалова С.А. Оценка здоровья школьников с использованием программного комплекса для многофакторного анализа и оценки показателей соматического и психоневрологического статуса / Сборник материалов XXV Международной научно-технической конференции. 2019. С. 959-963.
8. Чекалова С.А., Чекалова Н.Г., Силкин Ю.Р., Миронова А.И., Кожевникова Т.М., Азова Е.А., Азов Н.А. Функциональные резервы организма – ведущий критерий интегральной оценки здоровья учащихся / В книге: VI Балтийский конгресс по детской неврологии. сборник тезисов конгресса. 2016. С. 378-380.
9. Язык программирования python 3.7 [Электронный ресурс]. – Режим доступа: <https://www.python.org/downloads/release/python-370>
10. Py2exe, python библиотека [Электронный ресурс]. – Режим доступа: <http://www.py2exe.org>

E.V. Posypkina, S.A. Chekalova

OPTIMIZATION OF INFORMATION PROCESSING IN THE CHILDREN'S POLYCLINIC

*Nizhny Novgorod State Technical University n.a. R.E. Alekseev,
Privolzhsky Research Medical University*

The report describes a software package for the optimization of information processing process in a children's polyclinic. The developed software is aimed at increasing the efficiency of processing the received information and creating a database for analysis.

Keywords: database, software package, children's polyclinic, information

РЕШЕНИЕ ЗАДАЧ ОТДЕЛЕНИЯ ОБЪЕКТОВ В ПРИСУТСТВИИ НОСИТЕЛЯ В ПАКЕТЕ ПРОГРАММ ЛОГОС

1) ФГУП «РФЯЦ-ВНИИЭФ», ПАО «Компания «Сухой», г. Москва.

Доклад посвящён вопросам численного моделирования задач аэродинамики, осложнённых процессами движения обтекаемых объектов.

Ключевые слова: летательный аппарат, технология Химера, интерфейс, интерполяционный шаблон, пакет программ ЛОГОС.

Введение

Моделирование процессов газовой динамики представляет, как научный интерес в целях изучения физических явлений, так и практический интерес при проектировании новых образцов техники. В настоящее время в инженерной практике повышенное внимание уделяется особенностям работы техники на переходных режимах, обусловленных движением составных частей образца. Примером может служить определение АДХ летательного аппарата в процессе отделения груза и дальнейшая его динамика.

Решение задач отделения грузов в пакете программ ЛОГОС [1] реализуется, используя технологию расчёта на сетках с перекрытиями [2]. В случае моделирования на сетках с перекрытиями, получивших в научных кругах второе название – сетки типа Химера, расчёт производится на двух и более топологически несвязанных между собой дискретных моделях. Вычислительный процесс ведётся параллельно на всех сетках с учётом их взаимного расположения и взаимодействия.

В качестве иллюстрации возможностей разработанных вычислительных схем и алгоритмов, реализованных в пакете программ ЛОГОС, в работе представлено решение характерной задачи авиационной промышленности.

Технология расчёта на сетках с перекрытиями

Технология расчёта на сетках с перекрытиями предполагает генерацию независимых дискретных моделей для различных объектов, которые в дальнейшем объединяются в одну сетку с перекрытиями. Вычислительный процесс на основе технологии Химера учитывает наличие перекрытий и обеспечивает взаимодействие разделённых сеток. На рис.1 представлен пример сетки с перекрытиями вблизи профиля крыла.

Центральной задачей методики является сопряжение несвязанных доменов и построение интерполяционного шаблона для взаимодействия сеток. Ниже приведены определения, принятые в большинстве работ по методам расчёта на сетках типа Химера:

- интерфейс – открытая граница дискретной модели, значение полей на которой формируется на основании данных из окрестных ячеек;
- акцептор – грань интерфейса;
- доноры – ячейки, значение полей которых используются для восстановления информации на интерфейсах.

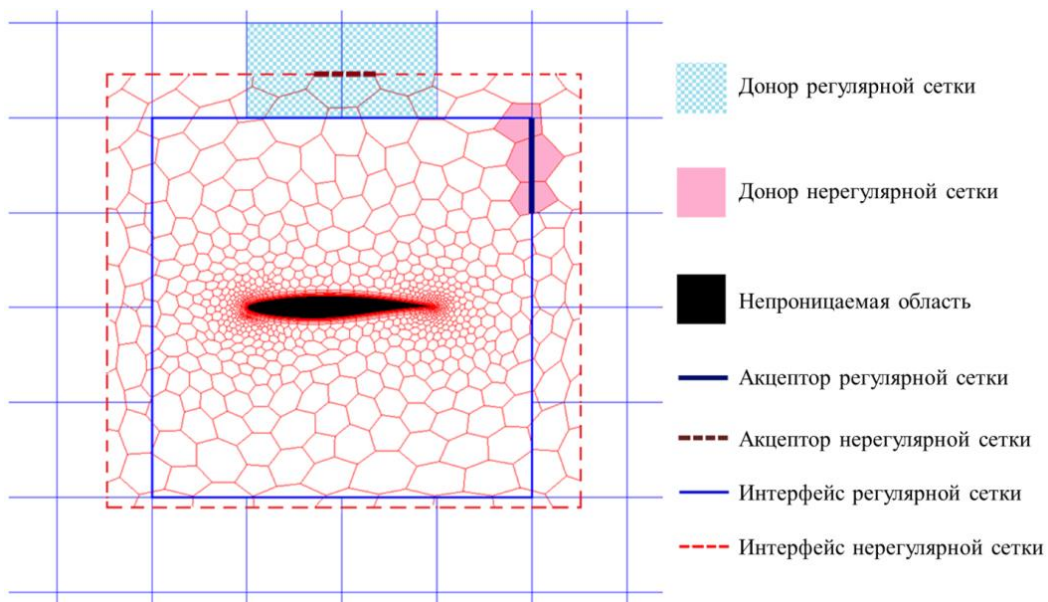


Рис. 1. Пример сетки с перекрытиями

Расчёт динамики отделения объекта

Отсоединяемый груз обладает своими динамическими свойствами, которые могут изменяться в процессе моделирования. В тех случаях, когда свойства объекта не изменяются во времени, целесообразно применять модельное понятие классической механики – модель твёрдого тела. Данная математическая абстракция позволяет проводить численный анализ движения тела под действием различного рода факторов, как под действием аэродинамических нагрузок, так и наличием сил гравитации. Для возможности численного моделирования подобных задач в пакете программ ЛОГОС внедрена модель твёрдого тела.

Ниже, на рис.2, приводится схема, описывающая последовательность этапов методики расчёта на сетках с перекрытиями с использованием модели твёрдого тела.



Рис. 2. Последовательность шагов

Динамика неполного подвесного топливного бака вследствие сброса будет отличаться от пустого, поскольку движение топлива в баке изменяет центр масс, инерционные характеристики объекта. В данном случае модель твёрдого тела неприменима, в качестве альтернативы может рассматриваться подход связанного расчёта: ЛОГОС-АЭРОДИНАМИКА+ЛОГОС-ГИДРОДИНАМИКА.

Расчёт динамики отделения объекта

В качестве демонстрации работоспособности разработанных алгоритмов рассматривается задача нестационарного отделения подвесного груза в трансзвуковом потоке вязкого газа [3] (рис. 3).

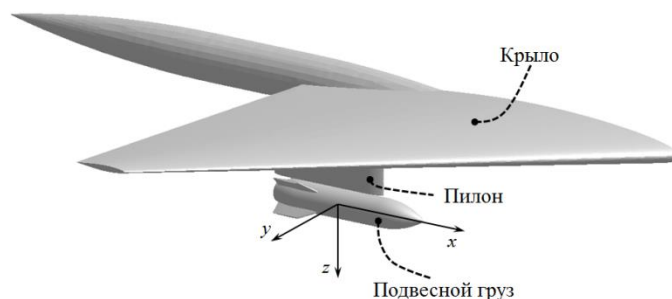


Рис. 3. Расчётная модель

Для решения задачи была сгенерирована сетка с перекрытиями, состоящая из 2 доменов. На груз на протяжении всего численного эксперимента задачи действовали аэродинамические нагрузки и сила тяжести; задавались масса, компоненты тензора инерции:

$$m = 907.1803 \text{ кг}, M_{xx} = 27.12 \text{ м}^2 \cdot \text{кг}, M_{yy} = M_{zz} = 488.1 \text{ м}^2 \cdot \text{кг}.$$

Для моделирования принимался однородный поток вязкого газа с числом Маха $M_\infty = 0.95$ при значении числа Рейнольдса $Re = 1.8 \times 10^6$. Шаг по времени равнялся 0.002 секунды. В задаче оценивается положения центр масс и изменение углов поворота груза с течением времени.

На рис. 4, 5 представлено сравнение результатов компьютерного моделирования и экспериментальных данных.

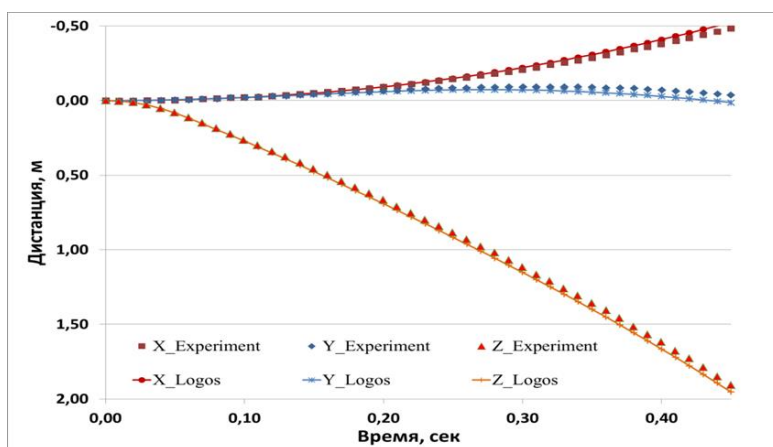


Рис. 4. Изменение координат центра масс в зависимости от времени

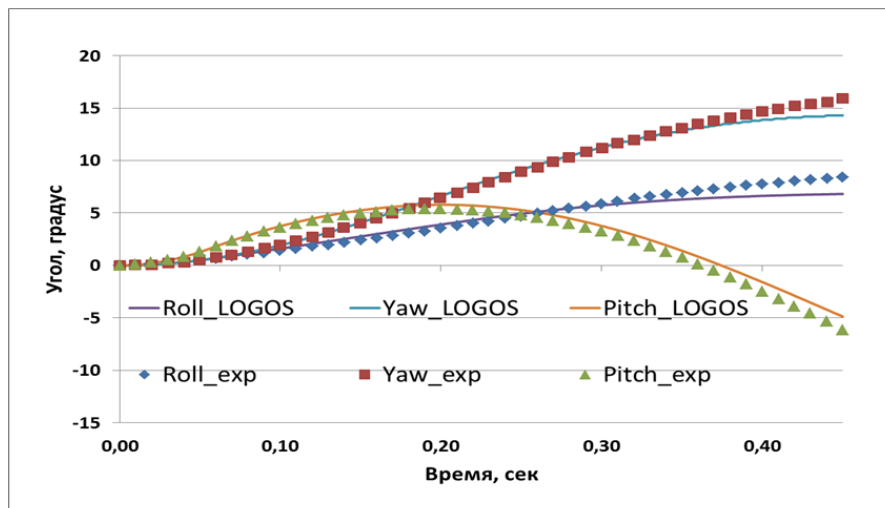


Рис. 5. Изменение углов курса, крена, тангажа в зависимости от времени

Как видно из представленных графиков, результаты расчёта хорошо согласуются с экспериментом.

Заключение

Пакет программ ЛОГОС позволяет успешно проводить численное моделирование процесса отделения грузов в присутствии носителя. Реализованные алгоритмы отвечают требованиям, предъявляемых к точности получаемых результатов, вычислительной эффективности.

Промышленная эксплуатация пакета программ ЛОГОС широко проводится в РФЯЦ–ВНИИЭФ и успешно применяется на предприятии ОКБ Сухого в процессе разработки новых образцов авиационной техники.

Библиографический список

1. Погосян М.А., Савельевских Е.П., Шагалиев Р.М., Козелков А.С., Стрелец Д.Ю., Рябов А.А., Корнев А.В., Дерюгин Ю.Н., Спиридонов В.Ф., Циберев К.В. Применение отечественных суперкомпьютерных технологий для создания перспективных образцов авиационной техники // Журнал ВАИТ, сер. Математическое моделирование физических процессов. – 2013. Вып.2. С. 3-17.
2. Дерюгин Ю.Н., Саразов А.В., Жучков Р.Н.. Особенности построения методики расчёта на сетках типа «Химера» для неструктурированных сеток // Математическое моделирование. – 2017. Т. 29, № 2. – С. 106-118.
3. Fox J. H., 23. Generic Wing, Pylon, and Moving Finned Store, Verification and Validation Data for Computational Unsteady Aerodynamics, RTO-TR-26, October 2000, St. Joseph Ottawa/Hill, Canada.

A.V. Sarazov¹, A.S. Kozelkov¹, A.V. Kornev², D.A. Ostanko²

SIMULATIONS OF SEPARATING OBJECTS IN THE PRESENCE OF THEIR CARRIER IN THE LOGOS SOFTWARE PACKAGE

¹FSUE “Russian Federal Nuclear Center – All-Russian Research Institute of Experimental Physics”, ²PJSC “Company” Sukhoi”, Moscow.

The paper deals with some aspects of computational aerodynamics complicated by flows over moving objects.

Keywords: flight vehicle, Chimera technology, interface, interpolation template, LOGOS software package.

УДК 512:53.08

Л.В. Грунская¹, В.В. Исакевич², Д.В. Исакевич²

ОБ ИСПОЛЬЗОВАНИИ АЙГЕНОСКОПИИ В ЗАДАЧАХ ОБРАБОТКИ, ХРАНЕНИЯ И ОТОБРАЖЕНИЯ ИНФОРМАЦИИ

¹Владимирский государственный университет имени А.Г. и Н.Г. Столетовых,

²ООО «Собственный вектор»

Подведены предварительные итоги 10-летнего развития новой информационной технологии — айгеноскопии, использующей для представления наблюдений адаптивные базисы собственных векторов ковариационных матриц конечномерных ансамблей временных сечений однородных наблюдений.

Предложенная технология обладает свойствами сверхчувствительности и сверхизбирательности по отношению к стандартным методам обработки, использующим фиксированные базисы, а также свойством предопределенности результата, позволяющим оперативно оценить выигрыш от использования айгеноскопии на начальных этапах разработки анализатора.

С использованием айгеноскопии обнаружено влияние гравитационных волн, излучаемых релятивистскими двойными звездными системами на инфранизкочастотное электрическое поле в приземном слое атмосферы, модернизирована система выявления реперных точек, предшествующих сильным землетрясениям на Камчатке по наблюдениям за почвенным радоном, создан прототип системы хранения и автоматического анализа сверхбольших массивов электрокардиограмм.

Ключевые слова: айгеноскопия, матрица ковариаций, сверхизбирательность, сверхчувствительность, собственный вектор, собственное значение, собственная пара, шумы неортогональности, гравитационная волна, землетрясение, электрокардиограмма.

Термин «айгеноскоп» впервые был введен в описании заявки на патент РФ № 116242 «Анализатор собственных векторов и компонент сигнала» [1]. После официальной публикации реферата на этот патент исчезли препятствия к дальнейшему использованию наименования «айгеноскоп» и «айгеноскопия». В настоящее время это вполне устоявшийся термин. Так, при запросе в поисковых системах по ключевому слову «айгеноскоп» находится несколько тысяч ссылок, ведущих к корневому использованию ссылки на патент № 116242.

Уже в момент подачи заявки на патент № 116242 было понятно, что введение в практику анализа сигналов айгеноскопии потребует обширной исследовательской программы, аналогичной той, которая была реализована по спектральному анализу перед выходом в свет книги А. Харкевича «Спектры и анализ».

Айгеноскоп разделяет сигнал на компоненты не так, как классический анализатор спектра. Например, синусоида, амплитудно модулированная синусоидой, представляется не тремя компонентами (несущая и две боковых), а шестью собственными векторами с кратностью собственных значений два и четыре. Первые два собственных вектора близки к гармоническим и представляют собой квадратуры несущей частоты, а остальные четыре — имеют характерную форму, неизменную вне зависимости от величины индекса модуляции, сохраняющуюся до исчезающе малых его значений.

Нами была реализована обширная программа исследования типологий (см., например, [2]) — поведения собственных пар (собственных векторов и спектров собственных

значений) в зависимости от того какие сигналы анализируются и как формируется из наблюдаемого сигнала ансамбль наблюдений. В большом числе случаев, представляющих практический интерес, поведение собственных векторов и спектров собственных значений позволяет судить по их виду о физической ситуации, порождающей наблюдаемый сигнал.

Что же порождает основные преимущества, которые реализуются айгеноскопами?

Стандартные методы анализа используют для представления сигналов фиксированные базисы — так, что структура компонент представления сигналов остается всегда неизменной — вне зависимости от структуры анализируемого сигнала. В этом «зародыш» (пока плохо осознанный специалистами) основного недостатка таких анализаторов — отклики каналов обработки таких анализаторов на полезный сигнал и на помеху имеют одинаковую форму и, следовательно, принципиально неотличимы в их смеси. Напротив, в айгеноскопе отклик на полезный сигнал индивидуален и может существенно отличаться от отклика на помеху, что открывает путь к сверхчувствительности и сверхизбирательности.

Представления в фиксированных базисах чрезвычайно эффективны, когда анализируемый сигнал действительно состоит из них, что характерно для систем связи и радиолокационных систем, работающих в линейных каналах и средах. Однако часто забывается, что само понятие «сигнал» изначально предполагает, что он имеет или искусственное происхождение, или, по крайней мере, имеется четкая договоренность о том, что считать сигналом.

В более сложных ситуациях (не сводящихся к вышеописанной) правильнее было бы говорить не о сигналах, а о (целеустремленных) наблюдениях за некоторыми временными сечениями окружающего мира, образующих упорядоченные числовые массивы (элементы многомерного векторного пространства). Айгеноскопия — вид целеустремленной деятельности (целеустремленных наблюдений), демонстрирующий свою эффективность именно в ситуациях (условиях), когда компоненты наблюдения обладают индивидуальной, априорно неизвестной структурой.

Подходы айгеноскопии вполне согласуются (а отчасти являются производными) от аксиоматики и теории, развитой в классической работе А. Акофа и Ф. Эмери «О целеустремленных системах» [3].

Единственное ограничение, которое накладывает айгеноскопия на свойства — они должны проявляться в некоррелированных компонентах наблюдений. Некоррелированность (независимость) компонент позволяет проявлять им, с одной стороны, достаточную индивидуальность, а с другой — проявлять эту индивидуальность независимо от других компонент. При этом структура этих компонент остается априори неизвестной. В такой индивидуальности нет ничего абсолютного, она представляет собой функциональное, а не структурное свойство и в такой же степени определяется наблюдателем, как и наблюдениями. При изменении параметров айгеноскопа структура компонент будет изменяться (до определенной степени); при этом компоненты будут сохранять индивидуальность своих свойств (например, спектральную или пространственную локализацию).

Айгеноскопия представляет собой вид целеустремленной деятельности по выявлению в ходе наблюдений определенных признаков, которые могут проявляться в виде определенных свойств некоррелированных с другими компонент наблюдений; структура компонент априорно неизвестна, а разумные предположения о свойствах таких компонент сформулированы (хотя бы на семантическом уровне).

На первом этапе айгеноскопии формируется ансамбль наблюдений — класс индивидуальных результатов наблюдений (объектов), то есть такое множество наблюдений, что замена любого из них любым другим наблюдением данного класса (в некотором окружении) не меняет индивидуальные свойства ансамбля.

На втором этапе айгеноскопии для ансамбля вычисляется матрица вторых смешанных моментов (матрица ковариаций, или ковариационная матрица, или — в

квантовой механике — матрица плотности), а для нее — на третьем этапе — вычисляются собственные пары (собственные векторы и собственные значения). Каждая собственная пара определяет свою (индивидуальную) аддитивную компоненту матрицы ковариаций; сумма этих компонент тождественно равна матрице ковариаций ансамбля.

Таким образом, если в ходе формирования ансамбля и выбора параметров наблюдений нам удастся сформировать ситуацию, в которой признаки будут индивидуально проявляться в виде отдельных «слоев ковариаций», то нам удастся их достаточно эффективно разделить на этапе анализа, и они полностью будут характеризоваться или одним, или несколькими собственными векторами (в том случае, когда собственные значения — кратные).

При айгеноскопии каждый из элементов ансамбля точно представим линейной комбинацией чистых (несмешиваемых) компонент (собственных векторов), а для своего приближенного (с любой заданной точностью) представления требует минимального числа компонент, пропорциональных собственным векторам.

Таким образом, собственные векторы и собственные значения полностью передают индивидуальные свойства ансамбля, хотя они априорно и неизвестны.

На последних этапах айгеноскопии индивидуальные свойства собственных пар ансамбля идентифицируются в соответствии с нашими представлениями о признаках, содержащихся (по нашему мнению) в ансамбле. На этапе идентификации ищется сходство между отдельными собственными векторами и, после их объединения в группы со сходными индивидуальными свойствами, ведется их сравнение с моделями признаков.

В айгеноскопии естественным образом объединяются целеустремленные действия, направленные на формирование групп наблюдений (ансамблей), в которых (по мнению наблюдателя) могут независимо друг от друга (некоррелированно) индивидуально проявляться признаки, которые гарантированно могут быть выявлены (вне зависимости от их индивидуальной структуры) в виде собственных векторов матрицы ковариаций и которые могут быть далее идентифицированы с моделями признаков.

Если при анализе не дать проявиться индивидуальной природе признаков, то признаки, лишённые индивидуальной природы, смешиваются на этапе «стандартного» анализа в фиксированном базисе, и этот эффект может быть фатальным.

Остановимся на важнейших свойствах айгеноскопии — сверхчувствительности и сверхизбирательности.

В основе явлений сверхчувствительности и сверхизбирательности лежит несмешиваемость некоррелированных компонент, присущая только представлению сигнала с использованием собственных векторов матрицы ковариаций. Поэтому свойства сверхчувствительности и сверхизбирательности связаны между собой. Остановимся на каждом из них отдельно.

Сверхизбирательность. Под избирательностью понимают способность приемника (анализатора) принимать отдельный сигнал без влияния других сигналов, которые при этом считаются мешающими. Если предполагается, что некоторая некоррелированная компонента, обладающая заранее известными свойствами (например, спектрально локализованная возле некоторой известной частоты), является сигналом или признаком, то под избирательностью айгеноскопа надо понимать его способность отделять такой сигнал от других (мешающих) некоррелированных компонент. Айгеноскоп делает это идеально — с точностью, определяемой шумами квантования и вычислительными погрешностями. При применении стандартных методов, основанных на использовании фиксированных базисов, некоррелированные компоненты оказываются неортогональными любому каналу приема, в том числе и тому, на котором ожидается появление признака. Возникают шумы неортогональности — избирательность падает (иногда весьма значительно). Таким образом, айгеноскоп всегда (за исключением искусственно созданных ситуаций и систем — см. выше) будет более избирателен, чем

стандартные анализаторы. При этом его избирательность будет ограничена только качеством технической реализации айгеноскопа. Это и есть сверхизбирательность (естественно, по отношению к анализаторам с фиксированным базисом).

Сверхчувствительность айгеноскопов, также, как и сверхизбирательность, является следствием того, что некоррелированные компоненты не смешиваются. Действительно, айгеноскоп, в котором некоррелированные компоненты не смешиваются, обнаружит любую из них в случае, если ее значение превышает аппаратные погрешности айгеноскопа. В классическом анализаторе с фиксированным базисом в любом из каналов анализа, кроме аппаратных, действуют принципиально неустраняемые шумы неортогональности, являющиеся следствием смешивания некоррелированных компонент. В этих условиях искомая компонента будет обнаружена только в том случае, если она заметно превышает уровень шумов ортогональности; это и определяет чувствительность классического анализатора, которая будет заведомо ниже, чем у айгеноскопа. То есть, по сравнению с классическим анализатором айгеноскоп обнаруживает свойство сверхчувствительности.

Сверхчувствительность и сверхизбирательность определяют важнейшее свойство айгеноскопии — **предопределенность выигрыша** по отношению к классическим анализаторам. Уже после выбора способа формирования ансамбля и вычисления матрицы ковариаций на предоставленных заказчиком ретроспективных данных становится возможным точно оценить проигрыш по отношению к айгеноскопии любого метода, использующего фиксированный базис. Таким образом, уже на этапе выбора способа формирования ансамбля становится понятным: есть ли смысл применять айгеноскопию, или можно ограничиться классическими методами обработки.

Основные преимущества айгеноскопии, важные для широкого класса применений, а именно сверхчувствительность и сверхизбирательность айгеноскопии, проиллюстрируем на примере открытия ранее неизвестного физического явления — эффекта воздействия релятивистских двойных звездных систем на электрическое поле в приземном слое атмосферы.

С помощью айгеноскопии были выявлены значимые аномалии поведения компонент напряженности электрического поля Земли на частотах гравитационных волн релятивистских двойных звездных систем (РДЗС), то есть было обнаружено воздействие гравитационных волн от РДЗС на глобальную электрическую цепь Земли. Эффект аномального поведения был обнаружен на частотах гравитационно-волнового излучения РДЗС из списка, составленного Джонстоном (<http://www.johnstonsarchive.net/relativity/binpulstable>).

В качестве исходного материала были использованы многолетние временные ряды наблюдений вертикальной составляющей электрического поля в приземном слое атмосферы, полученные на экспериментальном полигоне кафедры общей и прикладной физики Владимирского государственного университета и на станциях наблюдения Воейково, Верхнее Дуброво, Душети Росгидромета, предоставленные нам Я.М. Шварцем. Все рассмотренные временные ряды имели время дискретизации один час, длительность непрерывных наблюдений не менее 20 лет и достаточно высокую разрядность регистрирующего оборудования.

Исследования, проведенные с использованием классического анализатора амплитудного спектра, не позволили выявить какие-либо значимые аномалии поведения компонент напряженности электрического поля Земли на частотах гравитационного излучения РДЗС. Классический анализатор амплитудного спектра, настраиваемый на частоты гравитационного излучения РДЗС, не обнаруживает ничего — кроме шума (теперь интерпретируемые как «шумы неортогональности»).

Использование айгеноскопии позволило выявить некоррелированные составляющие, спектрально локализованные на частотах гравитационных волн, излучаемых РДЗС (см. иллюстрацию на рис. 1), но и убедительно доказать

статистическую значимость эффекта воздействия гравитационных волн от РДЗС на электрическое поле Земли. Полученные результаты подробно документированы в электронной монографии [4] и статье [5].

Подводя предварительные 10-летние итоги использования айгеноскопии при модернизации «стандартных» технологий отметим, что айгеноскопия подтвердила свою эффективность по крайней мере, для трех актуальных научно-технических задач, а именно:

Достоверное обнаружение влияния гравитационных волн, излучаемых релятивистскими двойными звездными системами, на инфранизкочастотное электрическое поле в приземном слое атмосферы, ставшее возможным вследствие сверхчувствительности и сверхизбирательности айгеноскопии [6].

Модернизация, на базе использования принципов айгеноскопии, системы выявления реперных точек, предшествующих сильным землетрясениям на Камчатке по наблюдениям за почвенным радоном (Камчатский филиал Геофизической службы РАН) [7,8].

Создание прототипа системы хранения и автоматического анализа сверхбольших массивов электрокардиограмм [9-14].

Решения всех трех вышеописанных задач хорошо документированы, осуществлены на базе специализированных айгеноскопов, программно реализованных на системе Scilab, а также на отечественной платформе «Эльбрус».

Эффективность предложенных решений основана на свойствах сверхчувствительности, сверхизбирательности айгеноскопии (по отношению к общепринятым способам обработки) и на свойстве предопределенности результата айгеноскопии, дающего дополнительное преимущество — возможность оценивать потенциальную эффективность айгеноскопии на самых ранних этапах проектирования.

Изложенное позволяет надеяться авторам, что айгеноскопия является действенным инструментом новых открытий и модернизации широкого класса технологий.

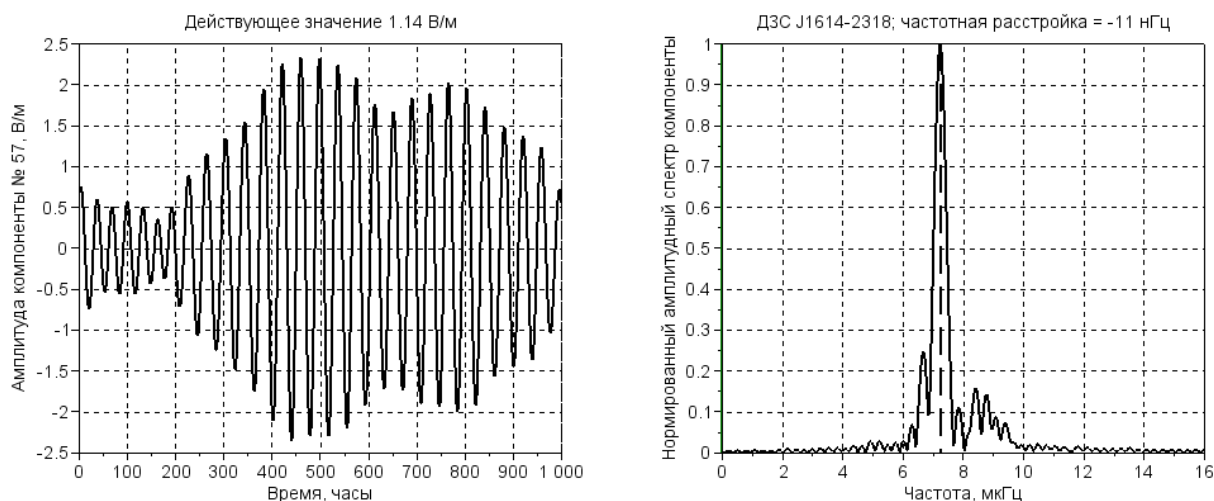


Рис. 1. Слева — некоррелированная компонента, соответствующая собственному вектору № 57 временного ряда станции Верхнее Дуброво на интервале анализа 1000 часов с действующим значением 1.14 В/м. Справа — нормированный к своему максимальному значению амплитудный спектр некоррелированной компоненты, представленной на левом графике. Амплитудный спектр компоненты локализован вблизи удвоенной частоты обращения РДЗС J1614-2318 (частота обращения 0.3616898×10^{-5} Гц) с расстройкой -11 нГц (вертикальный пунктир — удвоенная частота обращения РДЗС).

Выводы

Полученные результаты демонстрируют уже реализованные и потенциальные возможности айгеноскопии при модернизации широкого спектра известных технологий — в случае острой необходимости улучшения их рабочих характеристик.

Библиографический список

1. Исакевич В.В., Исакевич Д.В., Грунская Л.В. Анализатор собственных векторов и компонент сигнала. Патент РФ № 116242. Приоритет: от 30 сентября 2011 г.
2. Исакевич Д.В. Основы анализа собственных векторов и компонент регулярных колебаний. [Электронный ресурс] – М. Издательство Перо, 2015. – 140с. ISBN 978-5-00086-934-5
3. Акофф Р., Эмери Ф. О целеустремленных системах. Пер. С англ. Под ред. И.А. Ушакова. М.Ж Сов. Радио, 1974. 272 с.
4. Грунская Л.В., Исакевич В.В., Исакевич Д.В. Айгеноскопия многолетних временных рядов электромагнитного поля в приземном слое атмосферы крайне низкочастотного диапазона на частотах двойных звездных систем [Электронный ресурс] — М. Издательство Перо, 2016. – 276 с. ISBN 978-5-906927-50-7
5. Anomalous Behavior of the Terrestrial Electric Field Intensity at Multiple Frequencies of Relativistic Binary Star Systems. Gravitation and Cosmology, 2018, Vol 24, No 4, pp 384 – 392, ISSN 0202-2893
6. Исакевич В.В., Исакевич Д.В. Анализатор спектра собственных векторов и компонент сигнала. Патент РФ № 178399. Приоритет: от 28 июня 2017 г.
7. Исакевич В.В., Исакевич Д.В. Сигнализатор значимых отличий. Патент РФ № 133642. Приоритет: от 05 апреля 2013 г.
8. Фирстов П.П., Исакевич В.В., Макаров Е.О., Исакевич Д.В., Грунская Л.В. Применение методики айгеноскопии для поиска предвестников сильных землетрясений в поле почвенного радона (^{222}Rn) на Камчатке (август 2012 – август 2013 гг.). Сейсмические приборы, 2014, т.50, № 3, с. 63 – 75, ISSN 0131-6230
9. Исакевич В.В., Исакевич Д.В., Батин А.С. Кардиоайгеноскоп. Патент на РФ № 128470. Приоритет: от 09 августа 2012 г.
10. Исакевич В.В., Исакевич Д.В., Сушкова Л.Т., Аль-Барати Б.С. Устройство хранения и анализа ЭКГ. Патент РФ № 162110. Приоритет: от 23 июля 2015 г.
11. Исакевич В.В., Исакевич Д.В. Высокоточный одноканальный кардиоайгеноскоп с внутренней синхронизацией. Патент РФ № 177224. Приоритет: от 28 июня 2017 г.
12. Исакевич В.В., Исакевич Д.В. Высокоточный одноканальный кардиоайгеноскоп с внешней синхронизацией. Патент на полезную модель № 177963. Приоритет: от 28 июня 2017 г.
13. Исакевич В.В., Исакевич Д.В. Устройство хранения и автоматического анализа ЭКГ. Патент РФ № 180638. Приоритет: от 28 июня 2017 г.
14. Исакевич В.В., Исакевич Д.В. Высокоточный многоканальный кардиоайгеноскоп. Патент № 185708. Приоритет: от 31 мая 2018 г. РФ № 185708. РФ № 185708. Приоритет: от 31 мая 2018 г.

L.V. Grunskaya, V.V. Isakevich, D.V. Isakevich

**USING EIGENOSCOPY FOR YHT PROBLEMS
OF INFORMATION PROCESSING, STORING AND PRESENTING**

Vladimir State University, Eigenvector Ltd., Vladimir

The work summarizes the preliminary results of the ten-year-long development of eigenoscopy – the new information technology that presents the results of observations using the adaptive bases of covariance matrices eigenvectors for finite-length samples of simultaneous observations.

The presented technology is super-sensitive and super-selective compared to the standard processing methods that use fixed bases. It also has the property of result predetermination making it possible to estimate the benefit of using eigenoscopy at the very beginning of analyzer development.

Eigenoscopy has made it possible:

- to detect the impact of gravity waves from relativistic binary star systems on infra-low frequency terrestrial electric field;
- to improve the Kamchatka system of detecting reference points preceding strong earthquakes by surveying the soil radon;
- to create the prototype of system for storing and automatic analyzing super-big massives of electrocardiograms.

Keywords: eigenoscopy, covariance matrix, super-sensitivity, super-selectivity, eigenvector, eigenvalue, eigenpair, non-orthogonality noise, gravity wave, earthquake, electrocardiogram.

О СОБСТВЕННЫХ ВЕКТОРАХ И СОБСТВЕННЫХ ЗНАЧЕНИЯХ КОВАРИАЦИОННЫХ МАТРИЦ ТИПИЧНЫХ РАДИОТЕХНИЧЕСКИХ СИГНАЛОВ

*1Владимирский государственный университет имени А.Г. Столетова и Н.Г. Столетова
2ООО «Собственный вектор», г.Владимир*

Рассмотрены схемы анализа собственных векторов и спектров собственных значений ковариационных матриц сигналов: схема с скользящим интервалом анализа, схема с стробирующим интервалом анализа, схема с синхронным ансамблем. Показана применимость однопараметрического (в зависимости от периода стробирования) семейства ансамблей. Рассмотрены способы анализа собственных векторов и спектров собственных значений, оценки сложности представления сигнала. Показана применимость анализа собственных векторов для выявления: компонент гармонических, полигармонических (в том числе периодических и квазипериодических), в том числе с амплитудной модуляцией, линейным и квадратичным ростом, экспоненциальным нарастанием и затуханием, частотной и амплитудно-частотной модуляцией; параметрических колебаний, в том числе с большим количеством независимых движений; компонент импульсных комплексов. Показана важность синхронизации для упрощения представления сигнала. Показано преимущество анализа собственных векторов перед анализом сигнала в фиксированном базисе.

Ключевые слова: собственный вектор, собственное значение, ковариационная матрица, ансамбль, схема анализа, стробирование, синхронизация, параметрические колебания, ортонормированный базис, типология.

Введение

В радиотехнике известна схема анализа сигнала, основанная на его представлении в фиксированном ортонормированном базисе, определяемом целью анализа и решаемыми задачами. Среди известных базисов: естественный базис отсчётов сигнала, гармонический базис Фурье; базисы ортогональных полиномов и функций Эрмита, Лагерра, Лежандра; базисы функций Хаара, Радемахера, Уолша. Многие из базисов позволяют осуществлять представление сигнала не только на бесконечном, но и на конечном интервале анализа.

Все наблюдаемые радиотехнические сигналы конечны; бесконечные сигналы — лишь удобное приближение, эффективное во многих ситуациях.

Известны схемы и методы анализа и обработки сигнала, основанные на его представлении в адаптивном базисе собственных векторов ковариационной матрицы:

1. разложение Карунена-Лоэва, рассматривающее сигнал как ансамбль реализаций случайного процесса;
2. метод «Гусеница» (анализ сингулярного спектра), основанный на построении траекторной матрицы, строки которой — отрезки сигнала на скользящем интервале анализа;
3. метод главных компонент, уделяющий основное внимание представлению сигнала как множества векторов в многомерном векторном пространстве.

В перечисленных методах первостепенное внимание уделяется доминирующим — с относительно большими собственными значениями (и, соответственно, с относительно большими энергиями) — собственным векторам и компонентам сигналов. Однако в ряде фундаментальных и прикладных областей возникают задачи выявления (обнаружения и анализа) таких компонент сигналов, которые малы по энергии, но информативны на

конечных интервалах анализа. Выявление таких компонент требует особого анализа всей совокупности собственных векторов и собственных значений.

Для анализа собственных векторов и собственных значений (в том числе недоминирующих) авторами предложена конструкция «Анализатор собственных векторов и компонент сигналов» (патент на полезную модель № 116242). Эффективное применение предложенной конструкции требует решения следующих задач:

1. построения схем анализа сигналов, реализующих заявленную конструкцию и позволяющих выявлять представляющие интерес компоненты;
2. изучения поведения анализатора в зависимости от характера анализируемого сигнала и используемых схем анализа, а также обобщения и систематизации полученных при таком изучении сведений;
3. оценки потенциальных характеристик анализатора в сравнении с схемами, использующими стандартные базисы.

Схемы анализа

В соответствии с [1], конструкция анализатора собственных векторов и компонент сигнала состоит из четырёх блоков, соединённых последовательно:

1. блока масштабирования сигнала и аналого-цифрового преобразования;
2. блока вычисления ковариационной матрицы;
3. блока вычисления собственных векторов и собственных значений;
4. блока вычисления скалярных произведений и анализа признаков.

В блоке 1 осуществляется преобразование аналогового сигнала в цифровой, его масштабирование и предварительная обработка. Блок 2 вычисляет ковариационную матрицу, собственные векторы и собственные значения которой вычисляются в блоке 3. Блок 4 вычисляет главные компоненты — скалярные произведения собственных векторов и отрезков сигнала, подлежащих анализу — а также оценивает признаки собственных векторов на основании заданных критериев.

Вся совокупность отрезков сигнала, подлежащих анализу, образует ансамбль; отрезки сигнала, входящие в состав ансамбля, есть элементы этого ансамбля. Наличие ансамбля позволяет вычислить две матрицы: ковариационную матрицу — матрицу вторых начальных моментов между отсчётами на интервале анализа; матрицу Грама — матрицу скалярных произведений между отрезками сигнала.

Собственные векторы ковариационной матрицы позволяют вычислить главные компоненты — скалярные произведения этих векторов на элементы ансамбля. Известно, что главные компоненты представляют собой собственные векторы матрицы Грама [2].

В работах [3, 4] авторами были рассмотрены схемы анализа, оставляющие отрезки сигнала на интервалах анализа неизменными; эти схемы различаются между собой только расположением интервалов анализа на временной оси. Все рассмотренные схемы имеют на входе одномерный (одноканальный) сигнал.

По расположению интервалов анализа на временной оси различаются:

1. схема, использующая скользящий (с шагом один отсчёт) интервал анализа цифрового сигнала;
2. схема с стробированием, в которой интервал анализа смещается во времени с фиксированным шагом (периодом стробирования);
3. схема, формирующая синхронный ансамбль — совокупность интервалов анализа, синхронизированных с определённым классом событий.

Особого внимания заслуживает схема анализа, использующая стробирование. Эта схема позволяет вести анализ одного сигнала с разным периодом стробирования, тем самым предполагая использование не одного ансамбля отрезков сигнала, а целого семейства таких ансамблей. Как будет показано далее, семейство ансамблей, порождённое изменением одного параметра (периода стробирования), обнаруживает свойства, которые отсутствуют у какого-либо отдельного ансамбля и могут быть использованы в качестве

признаков, характеризующих сигнал. Эффективный анализ семейства ансамблей с использованием стробирования требует соблюдения двух условий, для которых необходимо расширить схему анализа:

1. все ансамбли семейства должны быть построены по одному и тому же сигналу — что требует ввести в схему анализа блок хранения сигнала (в аналоговом или цифровом виде);
2. шаг изменения частоты стробирования должен быть достаточно мал (существенно меньше частоты дискретизации) — что может потребовать добавления блока интерполяции цифрового сигнала.

Собственные векторы характеризуют форму выявляемых компонент сигнала, а собственные значения — среднюю энергию этих компонент на интервалах анализа.

Форма собственных векторов может быть охарактеризована визуально, с использованием критериев локализации или в аспекте общесистемных свойств. Визуальный анализ предполагает непосредственное участие исследователя и при большом размере ковариационной матрицы может быть весьма трудоёмок; тем не менее, визуальный анализ необходим при отсутствии априорных сведений о характере сигнала. Использование критериев локализации позволяет осуществлять быстрый отбор собственных векторов (или главных компонент), похожих по форме на искомый образец; в качестве образцов могут быть использованы:

1. для локализации вблизи искомой частоты — отрезки гармонического сигнала этой частоты;
2. для локализации вблизи интересующего момента времени — импульсные сигналы (векторы);
3. для выявления периодических сигналов — отрезки сигналов указанной формы (например, меандра) с произвольной начальной фазой.

К общесистемным [5] признакам собственных векторов, представляющим интерес при анализе, следует отнести их симметрию или асимметрию.

При отсутствии априорных сведений о характере сигнала собственные значения подлежат визуальному анализу. При использовании одного ансамбля анализу подлежит зависимость собственного значения (или нормированного — отнесённого к следу ковариационной матрицы — собственного значения) от его номера при упорядочении по убыванию; особое внимание уделяется наличию кратных собственных значений. При наличии семейства ансамблей, зависящего от одного параметра (в частности, от частоты стробирования), визуальному анализу подлежит зависимость собственных значений от параметра.

Интегральная оценка сложности представления сигнала в базисе собственных векторов, не требующая вычисления спектра собственных значений, основывается на отношении следа и нормы ковариационной матрицы; используемая норма вычисляется из суммы квадратов элементов матрицы [6].

Поведение собственных векторов и собственных значений

При анализе собственных векторов и собственных значений сложного сигнала с использованием какой-либо из перечисленных схем следует рассматривать сигнал как систему аддитивных компонент. Поведение собственных векторов сигнала определяется как свойствами (характером собственных векторов и видом спектра собственных значений) отдельных компонент, так и отношениями между ними (прежде всего, относительными энергиями компонент).

Гармонический (состоящий из одной гармонической компоненты) сигнал представляется наиболее просто — одним гармоническим собственным вектором, имеющим ненулевое собственное значение — при использовании стробирующего интервала анализа с периодом стробирования, кратным полупериоду сигнала. При другом периоде стробирования или при использовании скользящего интервала анализа наблюдаются два

гармонических собственных вектора — синфазный и квадратурный — имеющие близкие собственные значения.

Упрощение представления сигнала при кратности периода стробирования и полупериода сигнала обусловлено синхронизацией интервалов анализа и сигнала. Собственное значение единственного значащего собственного вектора гармонического сигнала при синхронизации интервалов анализа и сигнала превосходит каждое из собственных значений синфазного и квадратурного собственных векторов на 3 дБ.

Полигармонический сигнал представляется как совокупность нескольких компонент. Если синхронизация интервалов анализа и компоненты отсутствует, то компонента представлена парой гармонических (синфазных и квадратурных) собственных векторов с близкими собственными значениями. При наличии синхронизации наблюдается один гармонический собственный вектор.

Если полигармонический сигнал является периодическим, то все его гармонические компоненты характеризуются общим периодом, который совпадает с периодом сигнала. Поэтому использование стробирующих интервалов анализа с периодом стробирования, кратным периоду сигнала, позволяет существенно упростить представление сигнала — вплоть до одного периодического собственного вектора.

Если полигармонический сигнал является квазипериодическим — содержит гармонические компоненты с основными и комбинационными частотами — то его представление существенно упрощается при частотах стробирования, кратных этим частотам. Если две компоненты нарастают или затухают экспоненциально, то резкое изменение соответствующих собственных значений наблюдается при частотах стробирования, кратных разностной частоте компонент. Если одна из компонент затухает, а другая нарастает экспоненциально — то резкое изменение собственных значений наблюдается при частоте стробирования, кратной суммарной частоте компонент.

Компоненты амплитудно модулированного сигнала представляют собой компоненты несущей, амплитудно модулированные различными компонентами коэффициента. Например, гармонический сигнал с гармонической амплитудной модуляцией характеризуется гармоническими собственными векторами, имеющими соответствующую амплитудную модуляцию. При экспоненциальном нарастании или затухании сигнала собственные векторы приобретают экспоненциальное нарастание или затухание.

Линейное нарастание амплитуды гармонических колебаний приводит к появлению энергетически недоминирующих гармонических собственных векторов, амплитудно модулированных линейной функцией. Квадратичное нарастание — к появлению гармонических собственных векторов, амплитудно модулированных линейным и квадратичным полиномами Лежандра. Полиномы Лежандра в огибающих собственных векторов возникают потому, что они представляют собой систему полиномов, ортогональных на конечном интервале анализа с постоянным весовым коэффициентом.

Компоненты сигнала с монотонной частотной и амплитудно-частотной модуляцией соответствуют интервалам времени с разными значениями мгновенной частоты. Собственные значения (средние энергии) этих компонент определяются как длиной соответствующих интервалов — a , следовательно, и скоростью изменения мгновенной частоты — так и средней мощностью сигнала на этих интервалах. Вид спектра собственных значений определяется характером изменения мгновенной частоты и амплитуды колебаний во времени. В случае линейной частотной модуляции и использования скользящего интервала анализа частота и амплитуда собственных векторов возрастают вблизи краёв интервала анализа. При нелинейной частотной модуляции и использовании скользящего интервала анализа наблюдаются собственные векторы, имеющие асимметричную амплитудную и частотную модуляцию.

Сигналы, возникающие в линейных параметрических цепях, есть важный частный случай сигналов с амплитудной и частотной модуляцией. В простейшем случае (периодическое параметрическое воздействие на гармонический осциллятор) при периоде стробирования, кратном периоду параметрического воздействия, наблюдаются два гармонических собственных вектора с периодической амплитудной и частотной модуляцией; при других периодах стробирования синхронизация «рассыпается», и число значащих собственных векторов становится большим. Если в цепи имеет место параметрический резонанс, то также наблюдаются два собственных вектора: один из них с экспоненциальным нарастанием, а другой — с экспоненциальным затуханием.

Использование семейства ансамблей с разным периодом стробирования позволяет выявлять периодические параметрические воздействия на линейные цепи; как показывает вычислительный эксперимент [3], число «независимых движений» этих цепей может быть весьма велико.

Если сигнал представляет собой нерегулярную последовательность импульсных комплексов, то наиболее простое его представление обеспечивается при использовании синхронного ансамбля — то есть при синхронизации интервалов анализа с отдельными импульсными комплексами. Собственные векторы ковариационной матрицы синхронного ансамбля соответствуют компонентам импульсных комплексов. Ошибка синхронизации приводит к расширению спектра собственных значений и сглаживанию собственных векторов, представляющих компоненты импульсных комплексов.

Характеристики анализатора собственных векторов и компонент сигналов

В отличие от анализа сигнала в фиксированном базисе, анализ собственных векторов предполагает наличие априорно неизвестных признаков у самих базисных (собственных) векторов; для выявления этих признаков — как локализации, так и общесистемных — целесообразно использовать фиксированный базис.

В представлении сигнала в фиксированном базисе смешан вклад, обусловленный разными компонентами сигнала. Поэтому анализатор, обеспечивающий возможность индивидуального анализа собственных векторов ковариационной матрицы в фиксированном базисе, имеет потенциальное преимущество перед схемой, использующей фиксированный базис для представления сигнала [7]. Этот выигрыш ограничен ошибкой квантования сигнала и возрастает на 6 дБ/бит вместе с разрядностью аналого-цифрового преобразователя и на 10 дБ/декада вместе с объемом ансамбля, а также зависит от спектра собственных значений.

Библиографический список

1. Исакевич В.В., Исакевич Д.В., Грунская Л.В. Анализатор собственных векторов и компонент сигнала. Патент РФ № 116242. Приоритет: от 30 сентября 2011 г.
2. Главные компоненты временных рядов: метод «Гусеница». Под ред. Д.Л. Данилова и А.А. Жиглявского. СПбГУ, 1997. URL: www.gistatgroup.com/gus/book1/alg.html
3. Исакевич Д.В. Основы анализа собственных векторов и компонент регулярных колебаний. [Электронный ресурс] – М. Издательство Перо, 2015. – 140с. ISBN 978-5-00086-934-5.
4. Исакевич В.В., Исакевич Д.В., Батин А.С. Кардиоайгеноскоп. Патент на РФ № 128470. Приоритет: от 09 августа 2012 г.
5. Урманцев Ю.А. Общая теория систем в доступном изложении. – М.-Ижевск: НИЦ «Регулярная и хаотическая динамика», 2014. – 408 с.
6. Исакевич Д.В. Эффективная оценка размерности представления сигнала в базисе собственных векторов / М., «Радиотехника»: Успехи современной радиоэлектроники, №8, 2015, с.71-79.
7. Исакевич В.В., Исакевич Д.В. Анализатор спектра собственных векторов и компонент сигнала. Патент РФ № 178399. Приоритет: от 28 июня 2017 г.

L.V. Grunskaya, V.V. Isakevich, D.V. Isakevich

**ON EIGENVECTORS AND EIGENVALUES OF COVARIANCE MATRICES
OF TYPICAL SIGNALS**

Vladimir State University, Eigenvector Ltd., Vladimir

The covariance matrix eigenvectors' and eigenvectors' signal analysis schemes are considered which are the following: use of analysis span shift, use of span strobing, use of synchronous analysis span. The usefulness of parametric (depending from the strobing period) ensemble set is shown. The eigenvectors' and eigenvalues' spectra analysis are considered both with signal presentation complexity estimation. The usefulness of the eigenvectors' analysis for the detection of harmonic, polyharmonic (periodic, quasiperiodic etc.) components is shown including the components with amplitude and frequency modulation, exponential and polynomial growth and decay, parametric oscillations (also with high number of freedom degrees), pulse series' components. Importance of synchronization is shown for the effective signal' representation. The advantage of eigenvectors' analysis is shown.

Keywords: eigenvector, eigenvalue, covariance matrix, ensemble, analysis scheme, strobing, synchronization, parametric oscillations, orthonormal basis, typology.

**ВЫБОР ЗАКОНА СКАНИРОВАНИЯ ОБЛАСТИ ЦЕЛИ МАЛОГАБАРИТНЫМ
БОРТОВЫМ РАДИОЛОКАТОРОМ**

Филиал РФЯЦ-ВНИИЭФ «НИИИС им. Ю.Е. Седакова», г. Нижний Новгород

В докладе описываются преимущества и недостатки электронного и механического способов управления диаграммой направленности антенны бортового радиолокатора для осуществления сканирования области цели. Рассматривается возможность реализации механического способа сканирования в КВЧ диапазоне. Приводятся результаты математического моделирования наиболее распространенных законов сканирования: линейное построчное, линейное межстрочное, ортогональное спиральное, архимедово спиральное.

Ключевые слова: бортовой радиолокатор, закон сканирования, опорно-поворотное устройство, радиолокационное изображение, антенная система.

Бортовой радиолокатор обеспечивает излучение зондирующих и прием отраженных от области цели сигналов с последующим обнаружением в зондируемой области целеподобных объектов. Одной из важнейших задач при реализации сканирования области цели является обеспечение пространственной селекции передаваемых и принимаемых сигналов путем формирования необходимой диаграммы направленности антенны радиолокатора, а также управление ее диаграммой направленности в заданном диапазоне углов. Для обеспечения минимального времени обзора заданной области пространства наиболее предпочтительным является электронный способ сканирования [1], имеющий наибольшую скорость обзора. Быстродействие электронного сканирования ограничивается, в основном, инерционностью электрической схемы, обусловленной постоянными времени электрических цепей, причём эта инерция на несколько порядков меньше механической.

На сегодняшний день в бортовых радиолокаторах широко применяются активные фазированные антенные решетки (АФАР) сантиметрового диапазона длин волн, обеспечивающие широкоугольное электронное сканирование с обзором пространства более полусферы [2]. Но все возрастающие требования к уменьшению габаритов радиолокаторов с сохранением их разрешающей способности вынуждают смещаться в более коротковолновую область радиочастотного спектра – миллиметровый диапазон. Преимущества радиолокации в миллиметровом диапазоне длин волн в сравнении с сантиметровым и дециметровым диапазонами хорошо известны [3]: высокая разрешающая способность радиолокатора по угловым координатам, значительно меньшие габариты антенно-фидерных систем, допустимый уровень атмосферных и промышленных помех. В сравнении с оптическим и ИК диапазонами при работе в окнах прозрачности атмосферного поглощения в миллиметровом диапазоне обеспечивается большая дальность обнаружения в любое время суток как в простых, так и в сложных метеорологических условиях тумана, сильной запыленности или густой облачности. Однако, основной проблемой построения малогабаритных бортовых радиолокаторов с электронным управлением диаграммой направленности антенны, работающих в КВЧ диапазоне, является отсутствие отечественных малогабаритных серийных фазовращателей и активных элементов, подходящих для создания фазированных антенных решеток, поэтому при работе на миллиметровых волнах единственной возможностью обеспечения необходимого пространственного обзора области цели является реализация механического способа сканирования.

Механическое управление диаграммой направленности антенны (сканирующим лучом) в системе координат азимут-элевация может быть организовано по линейным, многолинейным, спиральным и т.д. законам благодаря использованию двухкоординатного опорно-поворотного устройства (ОПУ). Каждый закон сканирования имеет свои преимущества и недостатки, которые необходимо учитывать применительно к конкретным задачам. Для большинства бортовых радиолокаторов одним из ключевых параметров при выборе соответствующего закона сканирования является минимальное время получения полного кадра (радиолокационного изображения) в заданном секторе углов обзора. Рассмотрим наиболее распространенные законы сканирования:

- линейное построчное сканирование (рисунок 1а);
- линейное межстрочное сканирование (рисунок 1б);
- ортогональное спиральное сканирование (рисунок 1в);
- архимедово спиральное сканирование (рисунок 1г)

и проведем их математическое моделирование, где в качестве начальных условий для построения траектории движения сканирующего луча зададим:

- телесный сектор обзора в азимутально-угломестной системе координат: $\pm 30^\circ$;
- шаг сканирования: 1° ;
- фиксированное время перемещения между соседними координатными точками.

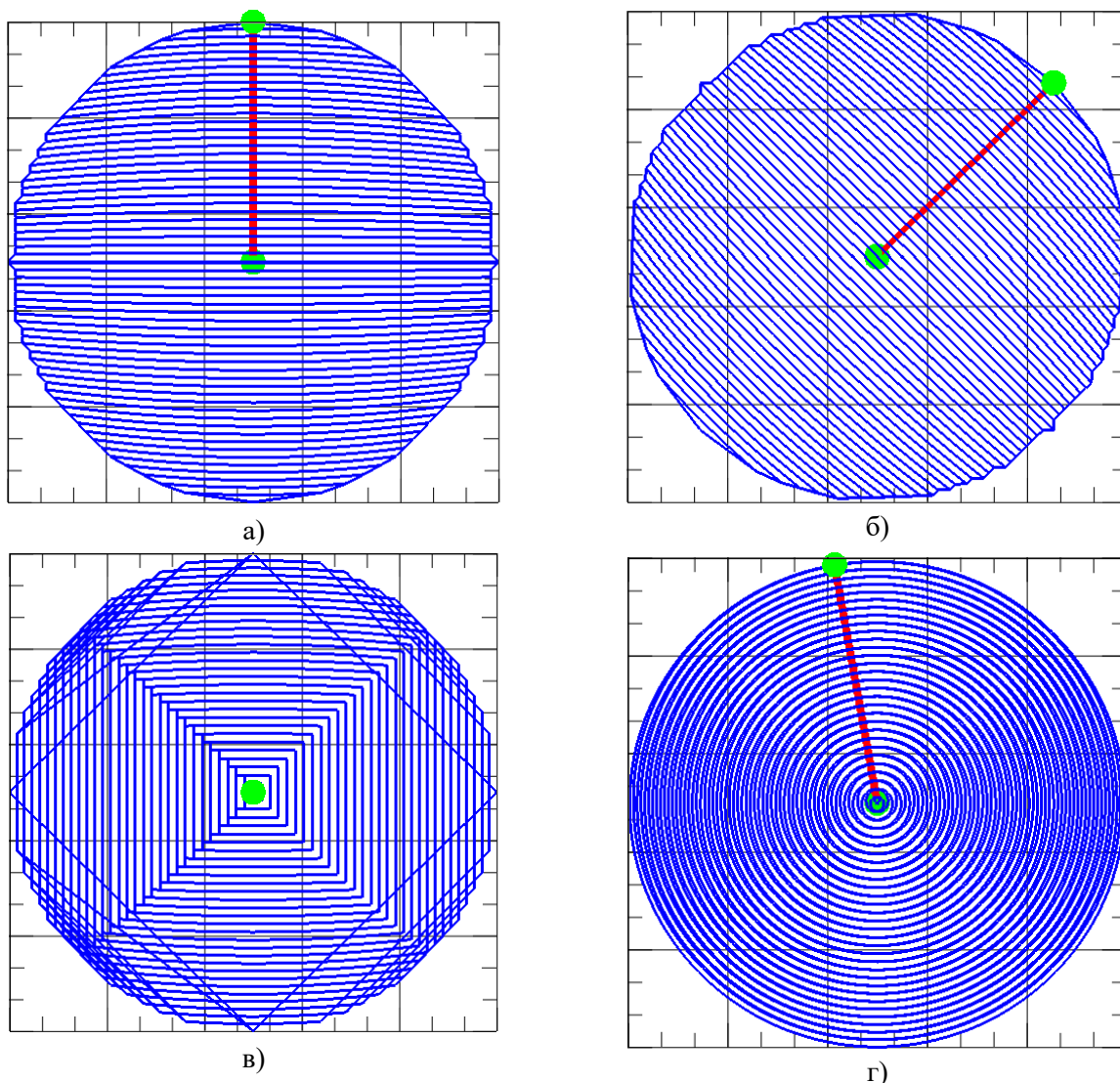


Рис. 1. Законы сканирования (траектории сканирующего луча):
 а) линейное построчное; б) линейное межстрочное;
 в) ортогональное спиральное; г) архимедово спиральное

В результате моделирования было выявлено, что при линейном построчном сканировании (рисунок 1а) основным рабочим механизмом является привод азимутальной оси. После отработки строки проводится смещение по углу элевации и реверс (смена направления вращения) по азимуту. После получения первого кадра (положение ОПУ – конечная точка) проводится реверс траектории движения. По завершению второго кадра ОПУ возвращается в начальную точку. Координаты приводов по азимуту и элевации приведены на рисунке 2 (здесь и далее сплошной линией обозначен график изменения угла по азимуту, пунктирной – по элевации).

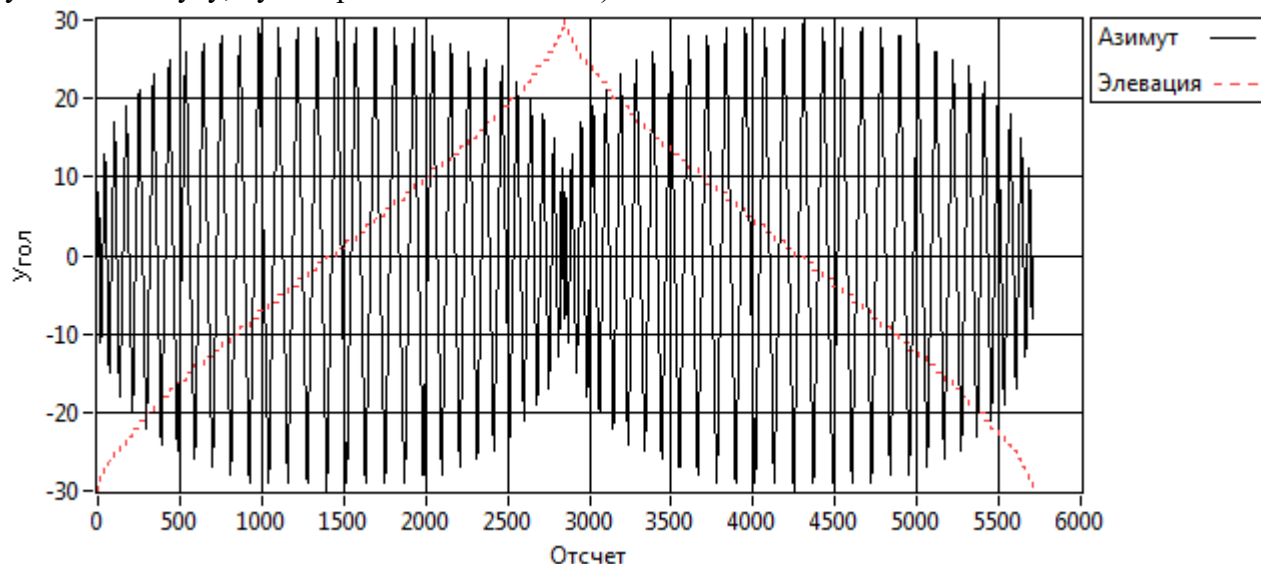


Рис. 2. Координаты приводов по азимуту и элевации линейного построчного сканирования

На основе анализа рисунка 2 было установлено, что для обеспечения средней относительной скорости порядка $1^\circ/\text{точку}$ при линейном построчном сканировании ускорение азимутальных приводов должно быть порядка $2\text{-}3^\circ/\text{точку}^2$, а при переходе к следующему кадру должно составлять $9^\circ/\text{точку}^2$.

При линейном межстрочном сканировании (рисунок 1б) наработка приводов становится равномерной (рисунок 3). Для обеспечения той же скорости $1^\circ/\text{точку}$ относительные ускорения снижаются до $1,5\text{-}2^\circ/\text{точку}^2$ в кадре и до $7^\circ/\text{точку}^2$ – при переходе к следующему кадру.

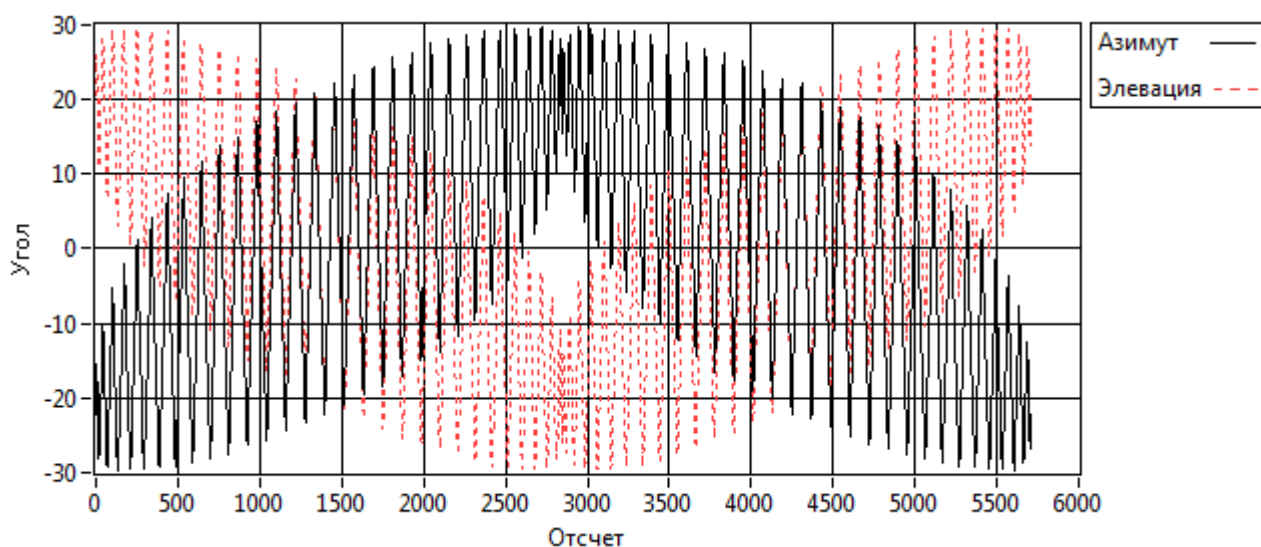


Рис. 3. Координаты приводов по азимуту и элевации линейного межстрочного сканирования

При ортогональном спиральном сканировании (рисунок 1в) начальным положением является точка с координатами азимут-элевация (0; 0). Первый кадр обеспечивается разворачивающейся ортогональной спиралью, второй – сворачивающейся. При ортогональном спиральном сканировании наработка приводов достаточно равномерная (рисунок 4), однако в зоне внешнего периметра траектории движения сканирующего луча, при условии сохранения телесного сектора обзора $\pm 30^\circ$, наблюдается резкий рост относительной скорости до $30^\circ/\text{точку}$ и относительного ускорения до $60^\circ/\text{точку}^2$.

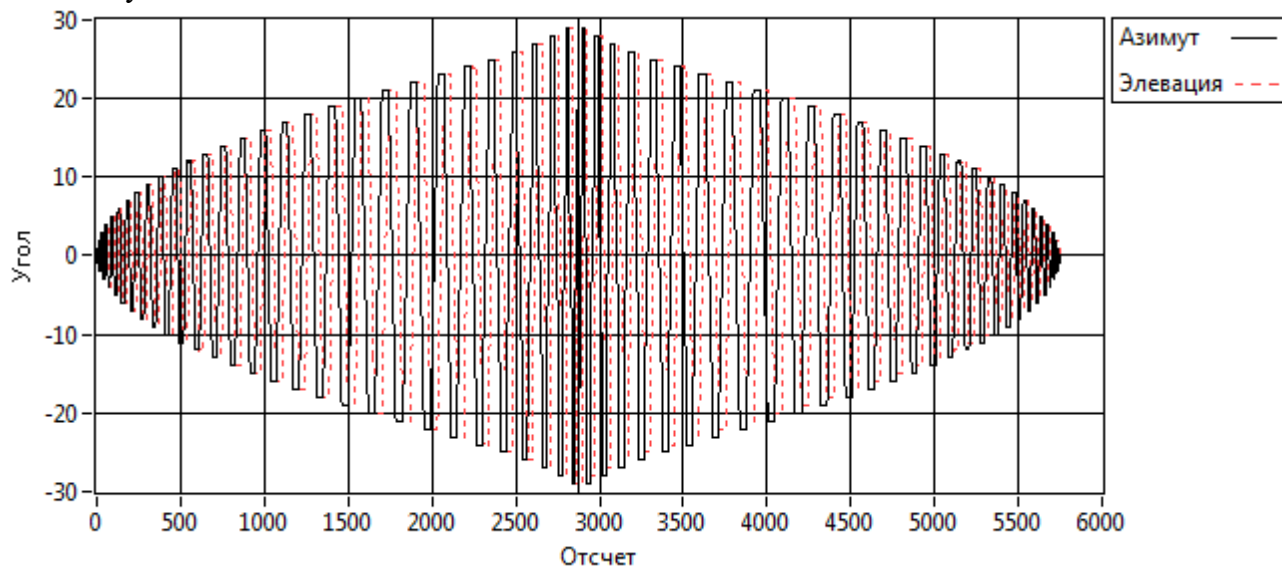


Рис. 4. Координаты приводов по азимуту и элевации ортогонального спирального сканирования

При всех вышеописанных способах механического управления положением антенной системы наблюдается целый ряд проблем, связанных с инерцией исполнительного механизма, что особенно проявляется при быстром сканировании. Частые смены направления движения увеличивают время сканирования вследствие процессов ускорения и торможения следящего привода. Устранить данный недостаток позволяет переход от кусочно-линейного закона изменения угловых координат приводов к гармоническому, так как при этом облегчается нагрузка на приводы, а траектория сканирующего луча приобретет вращательный характер.

На рисунке 1г приведена траектория спирального закона сканирования. Первый кадр – разворачивающаяся спираль Архимеда, второй кадр – сворачивающаяся. При этом внутри кадра смены направления движения не происходит. В данном случае при сохранении шага сканирования между витками спирали в 1° количество точек сканирования возрастает вдвое, по сравнению с предыдущими траекториями. Нарботка приводов азимут-элевация становится более равномерной (рисунок 5), а относительные скорости движения не превышают $1^\circ/\text{точку}$, при этом наблюдаются минимальные относительные ускорения – $0,03^\circ/\text{точку}^2$.

Таким образом, в результате использования архимедова спирального сканирования можно вдвое увеличить скорость движения приводов ОПУ, тем самым сократив время получения кадра в два раза, при этом относительные скорости и ускорения оказываются минимально возможными.

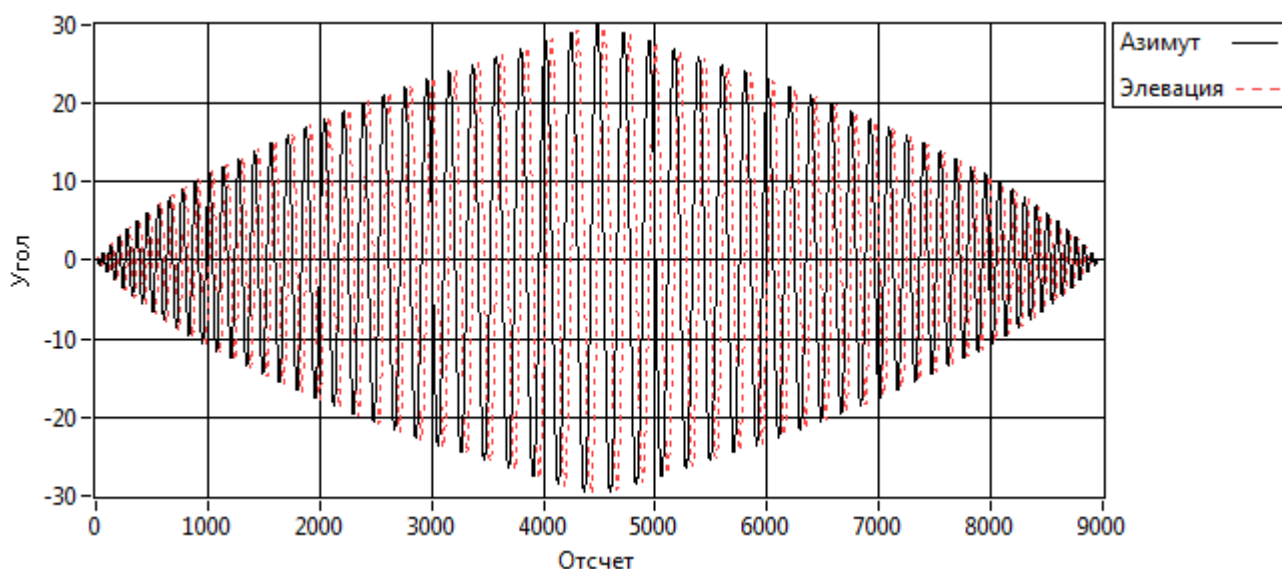


Рис. 5. Координаты приводов по азимуту и элевации архимедова спирального сканирования

С точки зрения снижения инерционности, уменьшения механической нагрузки, а, следовательно, снижения массогабаритных характеристик системы сканирования и повышения временной стабильности получения точек в кадре из рассмотренных законов сканирования архимедово спиральное сканирование является наиболее предпочтительным для применения в составе малогабаритного бортового радиолокатора при решении задач пространственного обзора области цели в миллиметровом диапазоне длин волн. Однако при использовании спирального закона сканирования алгоритм построения радиолокационного изображения, учитывающий изменение пространственной ориентации аппарата-носителя, становится более ресурсоемким ввиду невозможности представления результата без дополнительных преобразований в виде квадратной или прямоугольной матрицы с сохранением взаимного соответствия индексов точек и их координат. Это неизбежно приводит к увеличению времени на цифровую обработку по сравнению с другими рассмотренными случаями, а также требует значительного увеличения вычислительной мощности блока обработки радиолокатора.

Библиографический список

1. Скосырев В.Н. Методы обзора пространства в современных многофункциональных РЛС / В.Н. Скосырев // Наука и Образование. МГТУ им. Н.Э. Баумана. – 2014. – № 12. – С. 128–136.
2. Активные фазированные антенные решетки / Под ред. Д.И. Воскресенского и А.И. Канащенкова. – М.: Радиотехника, 2004. – 488 с.
3. Красюк В.Н. Особенности распространения радиоволн миллиметрового диапазона, перспективы их использования в современных радиотехнических системах / В.Н. Красюк, О.Ю. Платонов, А.Ю. Мельникова // Информационные каналы и среды. – №4. – 2003. – С. 33-38.

E.L. Varentsov, A.V. Golubtsov, N.A. Makarichev, K.V. Mineev, A.V. Nazarov

**SELECTION OF THE TARGET AREA SCANNING METHOD WITH SMALL-SIZED
ON-BOARD RADAR**

*Branch of the FSUE Russian Federal Nuclear Center – All-Russian Research Institute of
Experimental Physics Research Institute of Measuring Systems n.a. Yu.Ye. Sedakov,
Nizhny Novgorod, Russia*

The article describes the advantages and disadvantages of electronic and mechanical methods for controlling on-board radar antenna pattern for scanning the target area. The feasibility of mechanical scanning method in the EHF range is considered. The results of mathematical modeling of the most common scanning laws: linear, interlinear, orthogonal spirals, Archimedean spiral, are presented.

Keywords: on-board radar, scanning method, pivot device, radar image, antenna system.

МОДЕЛИРОВАНИЕ РАДИОЛОКАЦИОННОГО СИГНАЛА, ОТРАЖЕННОГО ОТ ПОДСТИЛАЮЩЕЙ ПОВЕРХНОСТИ

Филиал ФГУП РФЯЦ-ВНИИЭФ «НИИИС им. Ю.Е. Седякова»

Работа посвящена моделированию огибающей сигнала, отраженного от шероховатой подстилающей поверхности с использованием фацетной модели. Программно-реализованная модель учитывает изменение формы огибающей сигнала при отражении от подстилающей поверхности и его флуктуации при движении радиолокатора.

Ключевые слова: радиолокация, фацетная модель, огибающая сигнала, флуктуации отраженного сигнала.

При отработке алгоритмов работы широкого класса радиолокационных приборов возникает задача моделирования сигналов, отраженных от произвольной подстилающей поверхности. Для этого требуется разработка модели, обеспечивающей формирование огибающих тестовых сигналов. В настоящей работе приведено описание математической модели, позволяющей выполнить моделирование сигнала, отраженного от подстилающей поверхности, для радиолокационных приборов с импульсными зондирующими сигналами.

В настоящей работе для построения модели отраженного от подстилающей поверхности сигнала используется феноменологический подход [1], согласно которому протяженный радиолокационный объект или подстилающую поверхность можно представить в виде набора дискретных отражателей – фацетов, причем фазы элементарных сигналов рассеивателей взаимонезависимы. Тогда сигнал на входе приемной антенны равен сумме сигналов, отраженных от всех видимых из фазового центра АФС, статистически независимых фацетов.

При моделировании отраженного сигнала подстилающая поверхность задается в виде цифровой модели рельефа (карты высот) участка подстилающей поверхности, над которой выполняется моделирование. Для каждого элемента поверхности задается вид покрова подстилающей поверхности (вода, лес, пашня и т.д.) и соответствующие им отражательные характеристики (диаграммы обратного рассеяния). Для дальнейших расчетов рельеф представляется в виде набора фацетов треугольной формы, вершины которых находятся в узлах матрицы высот.

Для каждого фацета, находящегося в прямой видимости из фазового центра АФС, с помощью основного уравнения радиолокации рассчитывается принимаемый отраженный сигнал на входе приемного устройства радиолокатора. Видимость фацета из фазового центра антенны определяется с помощью алгоритма трассировки лучей, реализованного для минимизации времени моделирования, посредством параллельных вычислений с использованием библиотек NVIDIA CUDA.

Мощность P_i сигнала, отраженного от i -го фацета на входе приемного тракта равна:

$$P_i(t) = \frac{P_{\text{изл}}(t-\tau)G^2\sigma_i\lambda^2}{(4\pi)^3 R_i^4}, \quad (1)$$

где $P_i(t)$ и $P_{\text{изл}}(t)$ – зависимость мощности сигнала, принятого приемником и излученного передатчиком от времени соответственно, τ – задержка распространения сигнала, равная $\tau=2R_i/c$, G – коэффициент усиления антенны в направлении на фацет, σ_i – ЭПР фацета, рассчитываемая по формуле (2), λ – длина волны, R_i – расстояние между фазовым центром АФС и центром фацета.

ЭПР facets σ_i рассчитывается по формуле:

$$\sigma_i = \sigma_0(\theta) * S_i, \quad (2)$$

где $\sigma_0(\theta)$ – удельная ЭПР facets, зависящая от угла облучения (диаграмма обратного рассеяния), S_i – площадь facets.

Для получения сигнала $P_\Sigma(t)$, отраженного от всей подстилающей поверхности, выполняется суммирование отраженных сигналов от всех N facets, для которых выполняется условие прямой видимости:

$$P_\Sigma(t) = \sum_{i=1}^N P_i(t). \quad (3)$$

Результатом становится средняя (усредненная по ансамблю) форма отраженного сигнала.

Большинство реальных подстилающих поверхностей имеют неровности с высотой сравнимой или превышающей четверть длины волны зондирующего сигнала. Такие поверхности называются шероховатыми и для них характерно диффузное отражение, при котором когерентная компонента отсутствует. Радиосигнал, отраженный от шероховатой подстилающей поверхности, подвержен искажению формы [1].

В общем случае отраженный от земной поверхности сигнал подвержен двум видам флуктуаций: флуктуации от импульса к импульсу («медленные флуктуации») и флуктуации в пределах длительности принимаемого импульса («быстрые флуктуации»). Для получения достоверной модели отраженного сигнала необходим учет обоих видов флуктуаций.

Амплитуда отраженных от земной поверхности сигналов, принимаемых движущимся радиолокатором, подвержена интенсивным флуктуациям вследствие изменения сдвига фаз сигналов от различных участков облучаемой поверхности. Замирания сигналов рассмотренного вида обычно описываются доплеровским смещением частоты сигнала, обусловленным движением радиолокатора. Время, соответствующее исчезновению корреляции выборок, будет обратно пропорционально ширине доплеровского спектра принимаемого сигнала. Сигналы, соответствующие разным временным отсчетам АЦП на выходе амплитудного детектора, могут быть взаимно коррелированы. В различных сечениях эхоимпульса сигналы формируются различными участками поверхности. Когда интервал между сечениями менее длительности излучаемого импульса, удастся выявить общую часть поверхности, которая одновременно формирует сигналы в разных сечениях, что обуславливает взаимную зависимость сигналов. Учет внутритактовой и межтактовой корреляции производится скользящим весовым суммированием.

Согласно этому методу [2], из последовательности $x[n]$ с нулевым математическим ожиданием может быть получена последовательность $\psi[n]$ с заданной автокорреляционной функцией, с помощью скользящего суммирующего окна:

$$\begin{aligned} \psi[n] &= c_1 x[n-1] + \dots + c_N x[n-N], \\ \psi[n+1] &= c_1 x[n] + \dots + c_N x[n+1-N]. \end{aligned} \quad (4)$$

Коэффициенты $c_1 \dots c_N$ рассчитываются из системы уравнений

$$\begin{cases} R[0] = c_1^2 + \dots + c_N^2 \\ R[1] = c_1 c_2 + c_{N-1} c_N \\ \dots \\ R[N-1] = c_1 c_N \\ R[N] = 0. \end{cases} \quad (5)$$

Коэффициенты корреляции $R[0] \dots R[N]$ определяются, исходя из требуемой автокорреляционной функции. При моделировании быстрых флуктуаций коэффициенты корреляции, необходимые для расчета весовых коэффициентов окна, определяются автокорреляционной функцией зондирующего сигнала. Для моделирования медленных флуктуаций коэффициенты корреляции рассчитываются по доплеровскому спектру отраженного сигнала, который определяется вычислением доплеровского смещения для каждого facets и принимаемой мощности. Поскольку ширина доплеровского спектра изменяется в каждом сечении отраженного сигнала, его расчет производится в каждом временном срезе принятого импульса.

Огибающая принятого сигнала с учетом собственных шумов приемника может быть описана следующим выражением:

$$Q_i(t) = [A_i(t)y_{a,i}]^2 + [A_i(t)y_{b,i}]^2, \quad (6)$$

где $A_i(t)$ – средняя форма огибающей сигнала, i -ом сечении отраженного сигнала, $y_{a,i}$, $y_{b,i}$ – пара независимых нормальных псевдослучайных последовательностей (квадратурных компонент) принимаемого сигнала, каждая из которых является коррелированной последовательностью.

На рисунках 1...4 приведены результаты моделирования отраженного сигнала в дециметровом диапазоне длин волн для двух видов ровной однородной подстилающей поверхности: водная поверхность и заснеженный лес. При численном расчете использованы следующие значения параметров: высота приемопередатчика – 50 и 5000 м, длительность зондирующего импульса 0,1 и 4 мкс, мощность передатчика 1 Вт, ширина диаграммы направленности антенны 60° , максимум ДНА направлен в надир. Моделирование выполнялось как без учета флуктуаций, то есть моделировалась средняя форма отраженного сигнала, так и с их учетом. Пунктирной линией представлена средняя форма огибающей принятого сигнала, сплошной – одна из реализаций формы сигнала, рассчитанного с учетом флуктуаций. На рисунке 5 приведен результат моделирования сигнала, отраженного от каменистой подстилающей поверхности с развитым рельефом, представленным на рисунке 6. Длительность зондирующего импульса 1 мкс.

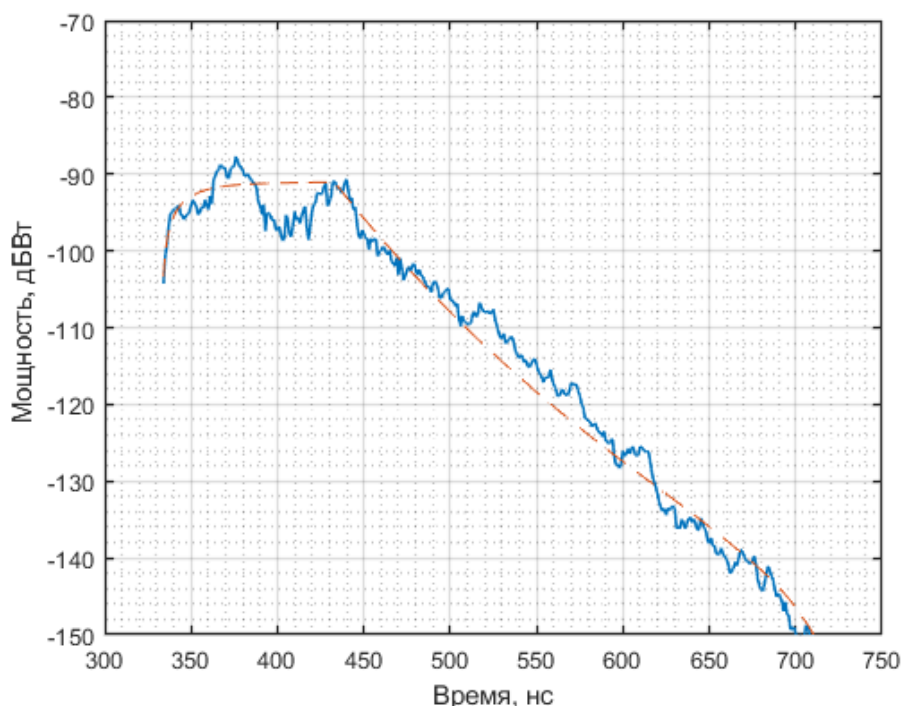


Рис. 1. Результат моделирования отраженного сигнала от поверхности типа «Заснеженный лес». Высота полета 50 м. Длительность зондирующего импульса 0,1 мкс

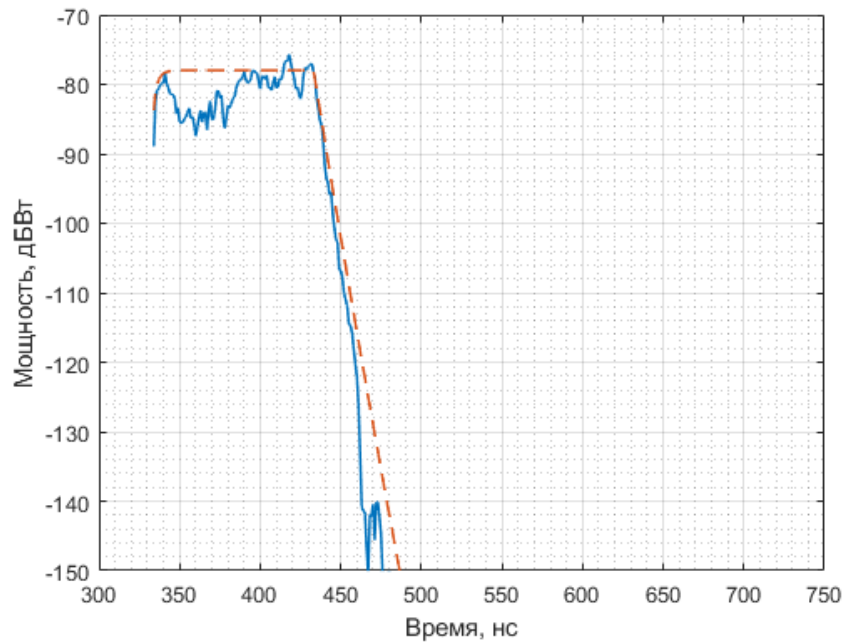


Рис. 2. Результат моделирования отраженного сигнала от поверхности типа «Водная поверхность с волнением 1-2 балла». Высота полета 50 м. Длительность зондирующего импульса 0,1 мкс

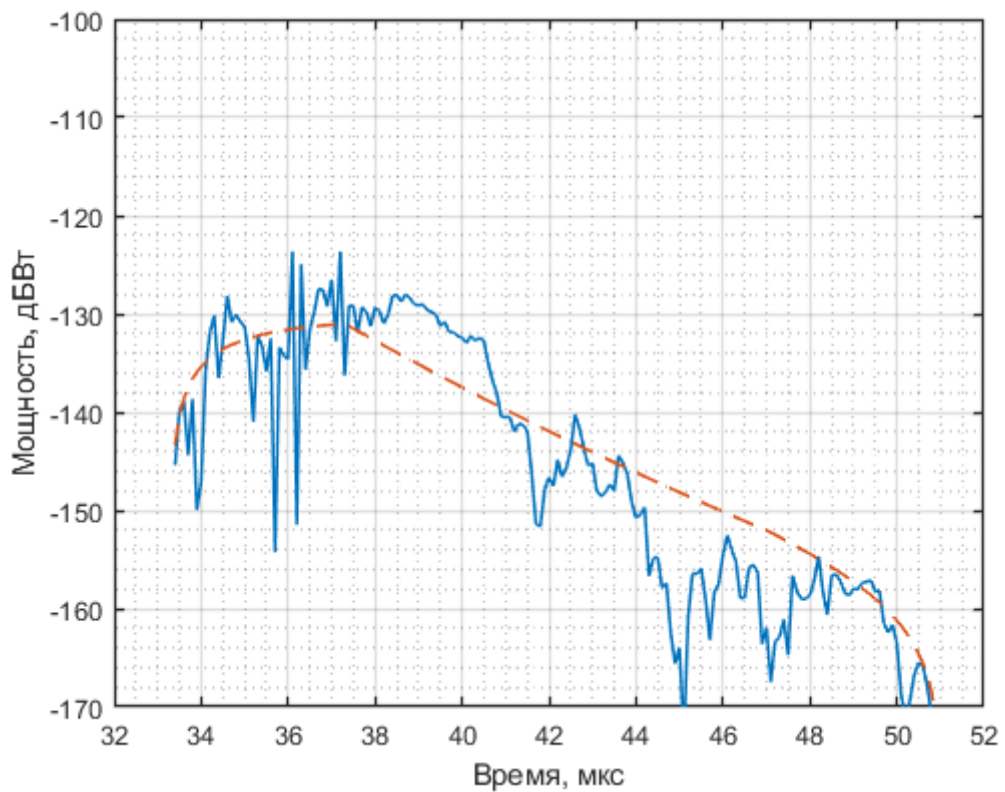


Рис. 3. Результат моделирования отраженного сигнала от поверхности типа «Заснеженный лес». Высота полета 5000 м. Длительность зондирующего импульса 4 мкс

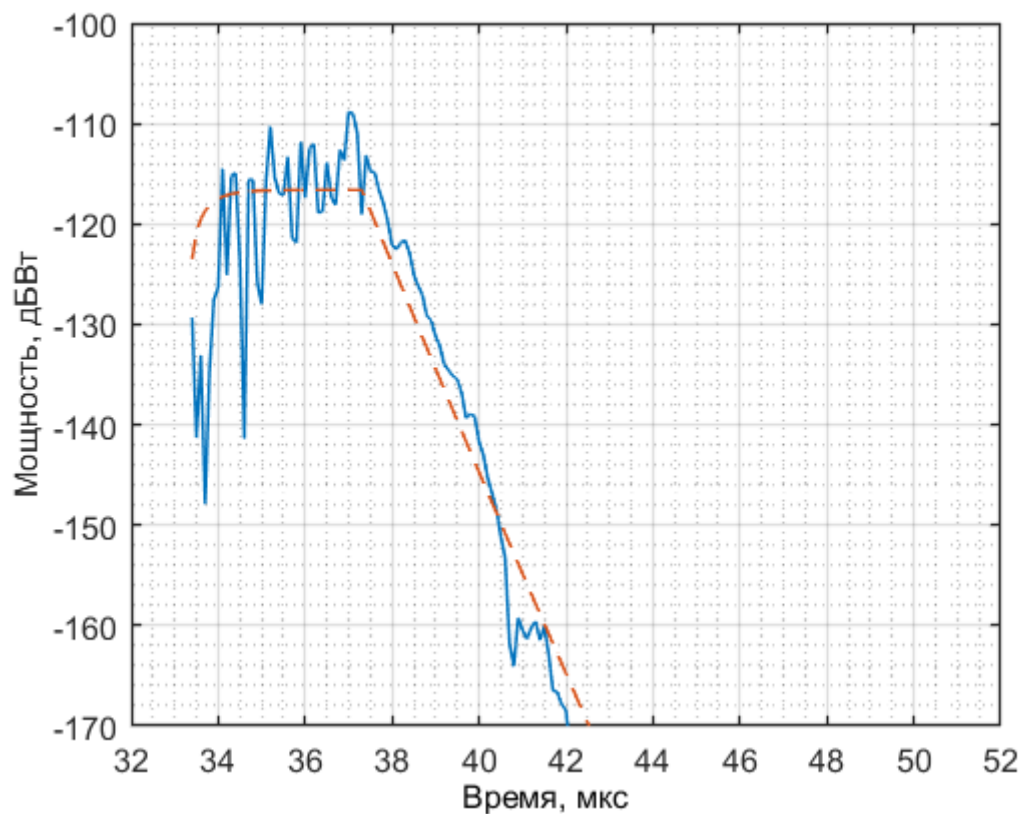


Рис. 4. Результат моделирования отраженного сигнала от поверхности типа «Водная поверхность с волнением 1-2 балла». Высота полета 5000 м. Длительность зондирующего импульса 4 мкс

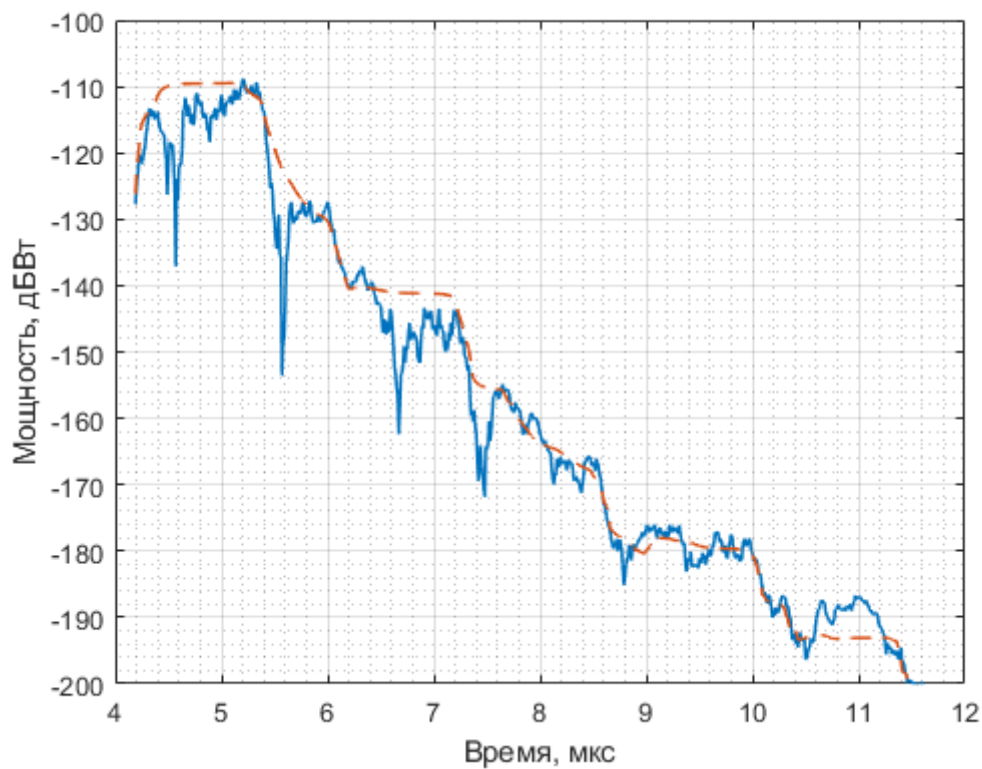


Рис. 5. Результат моделирования отраженного сигнала от поверхности типа «Заснеженный лес». Высота полета 1200 м. Длительность зондирующего импульса 1 мкс

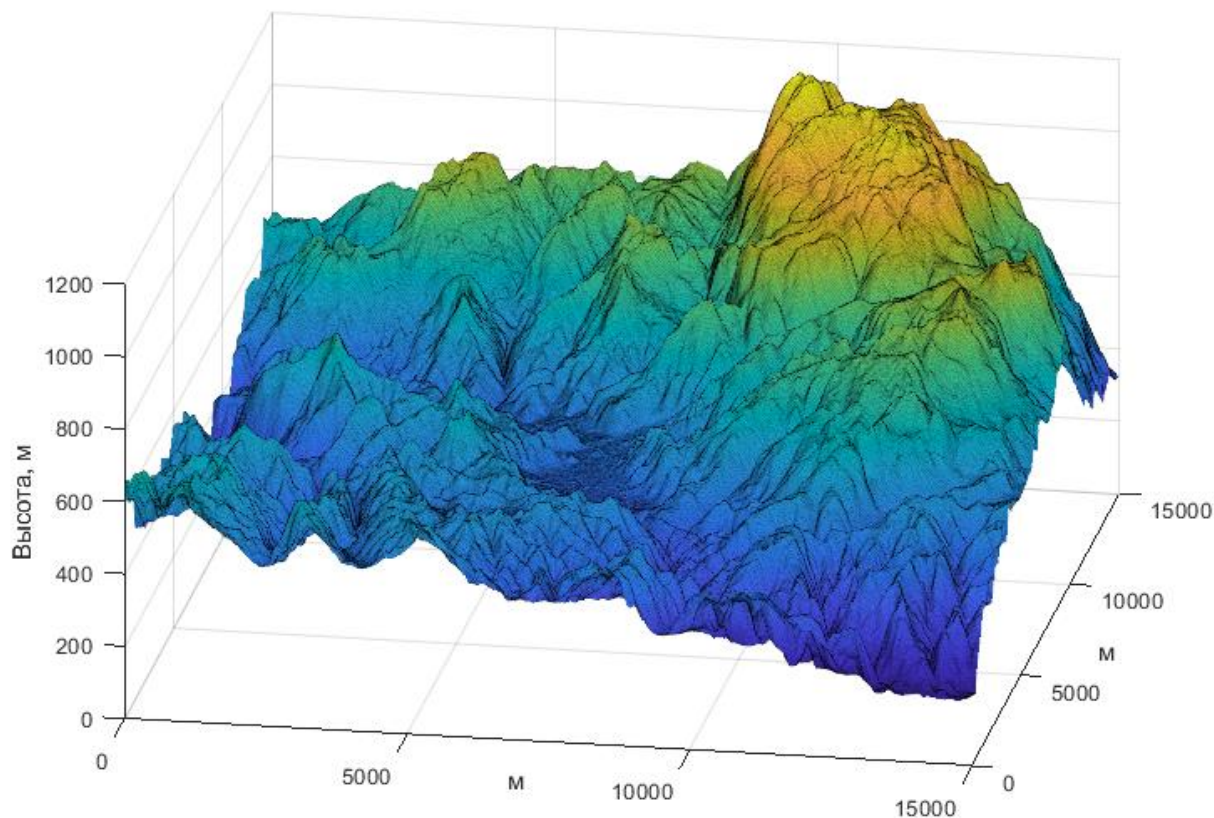


Рис. 6. Рельеф подстилающей поверхности

Из рисунков видно, что флуктуации оказывают значительное влияние на форму принимаемого сигнала. Полученные результаты согласуются с известными теоретическими результатами [1,3] и результатами экспериментальных измерений, полученными в НИИИС.

Библиографический список

1. Зубкович С. Г. Статистические характеристики радиосигналов, отраженных от земной поверхности / С. Г. Зубкович. - М.: Сов. радио, 1968. — 224 с.
2. Быков В.В. Цифровое моделирование в статистической радиотехнике/Сов. радио, 1971. – 328 с.
3. Жуковский А.П. Теоретические основы радиовысотометрии /А.П. Жуковский, Е.И. Оноприенко, В.И. Чижов. – М.: Сов. радио, 1979. – 320 с.

M.S. Dubrovin, Y.M. Kulikov, A.V. Anfimova

MODELING OF THE RADAR BACKSCATTER FROM A ROUGH SURFACES

Branch of Federal State Unitary Enterprise “Russian Federal Nuclear Center– All-Russian Scientific Research Institute of Experimental Physics” “Measuring Systems Research Institute n. a. Yu.Ye. Sedakov”

The paper deals with the digital modeling of echo signal envelope from rough surfaces with the Facet Model. Simulation results consider changes of signal envelope due to roughness variations and amplitude fluctuation caused by motion of the source.

Keywords: radar, facet model, signal envelope, amplitude fluctuation.

ТЕХНИЧЕСКИЙ ОБЛИК СТЕНДА КОНТРОЛЬНО-ПРОВЕРОЧНОЙ АППАРАТУРЫ ДЛЯ РАДИОЛОКАЦИОННЫХ ДАТЧИКОВ

ФГУП «РФЯЦ-ВНИИЭФ», г. Саров

В докладе описаны технический облик и варианты реализации узлов контрольно-проверочной аппаратуры (КПА) для регулировки и проверки характеристик радиолокационных датчиков.

Ключевые слова: радиолокационный датчик, система ближней радиолокации, контрольно-проверочная аппаратура, безэховая камера, линия задержки.

Радиолокационный датчик представляет собой радиолокационное устройство, включающее следующие основные функциональные части: приемопередающий модуль (ППМ), антенную систему (АС), устройство обработки сигнала, преобразователь напряжения питания.

Такие параметры РД, как дальность действия, разрешающая способность, чувствительность, универсальность, массогабаритные характеристики в значительной степени определяют эффективность всей системы или комплекса, в которую входит датчик.

Требования по дальности срабатывания оказывают доминирующее влияние на выбор принципа действия, состав и значения основных характеристик РД. При измерении малых дальностей реализация радиолокационных систем сопряжена с наличием возможного сигнала местной связи в приемнике от зондирующего сигнала. По этой причине применяется принцип работы с непрерывным излучением и линейной частотной модуляцией (ЛЧМ) зондирующего сигнала.

Определение дальности до цели методом ЛЧМ основано на измерении приращения несущей частоты передатчика за время распространения сигнала до цели и обратно.

Развитие систем ближней радиолокации, в частности РД, требует от разработчиков постоянного совершенствования характеристик разрабатываемых изделий [1].

Переход данных разработок в серийное производство требует оптимизации и автоматизации процедуры проверки параметров РД. Таким образом, задача разработки контрольно-проверочной аппаратуры (КПА) для первичного контроля характеристик больших партий РД миллиметрового и сантиметрового диапазона длин волн, является в настоящее время весьма актуальной.

Технический облик КПА должен быть разработан в соответствии с современными тенденциями развития промышленности, к которым относятся: минимизация использования импортного оборудования и комплектующих, в том числе и за счет собственного производства узлов, универсальность использования как самой КПА, так и её составляющих (требуется предусмотреть возможность использования данного оборудования не только для одного конкретного РД, но и для серии изделий).

Обобщенная структурная схема КПА представлена на рисунке 1.

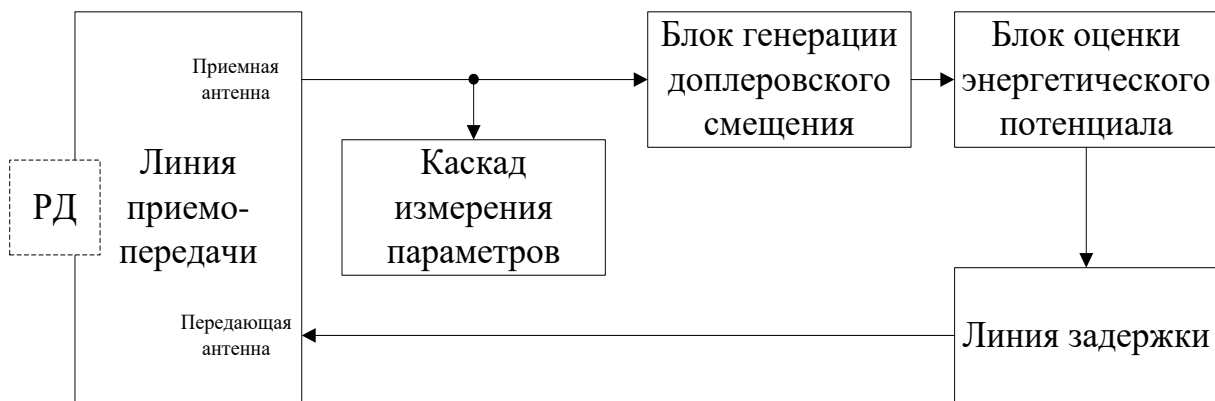


Рис. 1. Обобщенная структурная схема КПА

Линия приемо-передачи

Линия приемо-передачи (ЛПП) представляет собой устройство, которое обеспечивает прием и передачу сигнала от РД к узлам КПА и обратно без помех и искажений с помощью, передающей и приемной антенн.

В связи с тем, что РД, как правило, излучают сигнал с помощью АС, имеющей широкую диаграмму направленности (ДН) с максимумами коэффициента усиления (КУ), приходящимися на углы порядка $\pm (30-50)^\circ$, приемная и передающая антенны ЛПП должны располагаться в указанном диапазоне углов, что накладывает определенные требования на габаритные размеры ЛПП. Пример типового углового распределения КУ антенной системы РД представлен на рисунке 2.

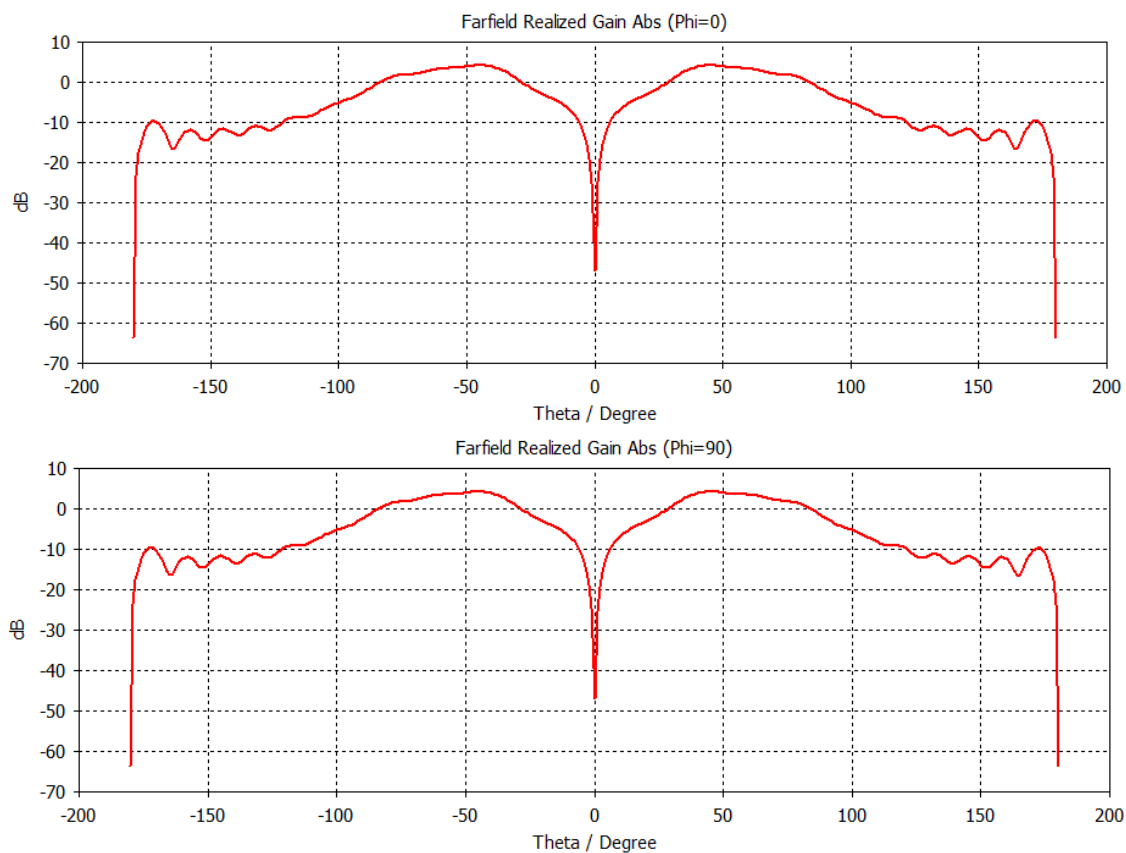


Рис. 2. Пример типового углового распределения КУ антенной системы РД в *E* и *H* плоскостях

При разработке КПА имеет смысл рассмотреть возможность реализации в ЛПП нескольких пар антенн для исследования характеристик РД при различных углах подхода, а также проверки РД, имеющих узкую ДН с максимумом КУ при угле 0° .

Так как длина ЛПП должна быть достаточна для формирования требуемой ДН (не менее десяти длин волн), то и ширина, позволяющая разместить приемную и передающую антенны на углах порядка $\pm (30-50)^\circ$, должна быть соответствующей. Для удобства исполнения приемную и передающие антенны следует разместить с разных сторон камеры. Таким образом, пространство ЛПП должно представлять собой параллелепипед с габаритами порядка нескольких десятков сантиметров. Также стоит рассмотреть вопрос использования одной ЛПП со сменными антенными системами для РД миллиметрового и сантиметрового диапазонов длин волн, поскольку зачастую размеры РД и габаритные размеры их обтекателей схожи.

ЛПП может быть реализована двумя способами: в виде кожуха из радиопоглощающего материала (РПМ) или безэховой камеры. Кожух, как правило, используется для РД с узкой ДН и отличается удобством крепления и простотой реализации. Кожух для РД с узкой ДН имеет ширину и высоту строго по габаритам самого изделия и надевается как «носок», что является крайне выгодным по сравнению с жесткой фиксацией в безэховой камере, особенно когда речь идет о проверке целой партии изделий. На конце кожуха размещается антенная система, обычно состоящая только из передающей антенны, установка двух антенн (приемной и передающей) может вызвать дополнительные искажения по причине недостаточной развязки, связанной с малым пространством кожуха. Ввиду больших требуемых размеров конструкция кожуха не сможет иметь достаточной жесткости или требует установки дополнительных ребер по всему объему кожуха, что нивелирует такое преимущество как простота реализации. Удобство крепления РД в кожухе также вызывает сомнения, так как для формирования требуемой ДН габариты кожуха должны быть значительными. Таким образом, на этапе формирования технического облика КПА наиболее перспективным вариантом реализации ЛПП является безэховая камера (примеры исполнения безэховой камеры представлены на рисунке 3[2,3]). При этом на этапе проектирования особое внимание следует уделить способу фиксации РД в камере с упором на простоту крепежа изделия.

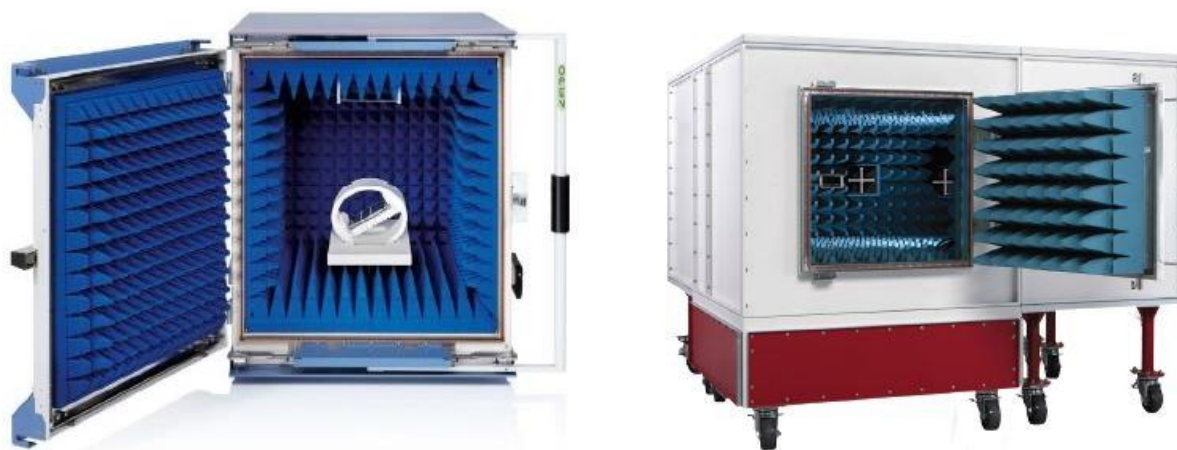


Рис. 3. Примеры исполнения безэховой камеры

Каскад измерения параметров

Каскад измерения параметров позволяет измерять характеристики излучаемого радиодатчиком сигнала до момента его преобразования. Измерению подлежат такие характеристики излучаемого сигнала как частота и мощность (с учетом ослабления в предшествующих узлах КПА: безэховой камере, соединителях, направленном ответвителе).

Блок генерации доплеровского сдвига

Блок генерации доплеровского сдвига предназначен для имитации смещения принимаемого сигнала, связанного с движением РД. Частота биений в статическом положении РД определяется по формуле [4]:

$$f_{\text{б}} = \frac{2R}{c} \cdot \frac{\Delta F}{\Delta T}, \quad (1)$$

где $f_{\text{б}}$ – частота биений (дальномерная частота); R – расстояние до ПП; c – скорость света; ΔF – девиация частоты; ΔT – период модуляции, и для типичных характеристик РД миллиметрового и сантиметрового диапазона длин волн составляет порядка единиц МГц.

Величина доплеровского смещения частоты определяется по формуле:

$$|F_{\text{д}}| = \frac{2V_{\text{макс}}}{\lambda_{\text{р}}}, \quad (2)$$

где $F_{\text{д}}$ – частота Доплера; $V_{\text{макс}}$ – максимальная радиальная скорость движения изделия; $\lambda_{\text{р}}$ – рабочая длина волны излучаемого сигнала.

При типичных скоростях движения изделия величина доплеровского смещения составляет порядка 100 кГц в миллиметровом и порядка 30 кГц – в сантиметровом диапазоне длин волн. При типичной дальности срабатывания, составляющей несколько метров, данное смещение создает отклонение не более двух десятков сантиметров. Таким образом, блок генерации доплеровского смещения является не обязательным узлом КПА, который должен быть реализован лишь в устройствах, предназначенных для проверки РД с малыми допусками по дальности срабатывания. Следует отметить, что блок генерации доплеровского сдвига дополнительно может быть использован для формирования помех.

Блок оценки энергетического потенциала

Блок оценки энергетического потенциала предлагается реализовать на базе переменного аттенюатора. Типичный энергетический потенциал РД составляет порядка 100 дБ. С учетом потерь в линиях КПА для миллиметрового диапазона длин волн достаточно применить аттенюатор, позволяющий изменять ослабление в пределах от 0 до 50 дБ, для сантиметрового диапазона (ввиду меньших потерь) следует использовать аттенюатор с более широким диапазоном ослаблений – до 100 дБ.

Линия задержки

Линия задержки (ЛЗ) предназначена для имитации пройденного сигналом расстояния до цели и обратно и позволяет контролировать дальность срабатывания РД. ЛЗ с преобразованием сигнала (например, волоконно-оптические ЛЗ) не имеют серийно-изготавливаемых образцов, делаются на заказ иностранными фирмами и являются достаточно дорогостоящими изделиями. Использование подобных изделий нецелесообразно при актуальной тенденции к импортозамещению. Поэтому предлагается реализация ЛЗ миллиметрового диапазона длин волн на системе металлических волноводов. Преимущество данного технического решения состоит в простоте изготовления и возможности регулировки длины линии путем соединения необходимого количества участков волноводов.

ЛЗ сантиметрового диапазона длин волн может быть реализована на коаксиальных кабелях. Недостатком данного технического решения является более сложная реализация возможности изменения длины ЛЗ.

Таким образом, предложен технический облик КПА для РД миллиметрового и сантиметрового диапазона длин волн, который позволяет одновременно контролировать частоту и мощность излучаемого РД сигнала, энергетический потенциал и дальность срабатывания. Преимуществом принятых технических решений является возможность изготовления основных узлов КПА в НИИИС им. Ю.Е. Седакова.

Библиографический список

1. Дорогов А.Д., Калинин Д.В., Минеев С.Е., Назаров А.В., Рыбаков Д.К., Приемопередающий модуль миллиметрового диапазона длин волн с повышенной стойкостью к ударным нагрузкам // Информационные системы и технологии. ИСТ – 2019: Материалы XXV Международной научно-технической конференции. – Н. Новгород, 2019. – С. 53-58.
2. https://www.rohde-schwarz.com/ru/product/dst200-productstartpage_63493-11087.html
3. <http://www.ets-lindgren.com/datasheet/test-systems/wireless-solutions/5009/500908>
4. Радиолокационные устройства (теория и принципы построения). Под ред. В.В. Григорина - Рябова. М.: Сов.радио, 1970. – С. 10.

A.D. Dorogov, S.V. Marevsky, A.V. Nazarov, R.R. Osmanov

**TECHNICAL APPEARANCE OF THE CONTROL AND TESTING EQUIPMENT
STAND FOR RADAR SENSORS**

*Federal State Unitary Enterprise “Russian Federal Nuclear Center– All-Russian
Scientific Research Institute of Experimental Physics”*

The article describes the technical appearance and implementation options for control and testing equipment designed to adjusting and checking the characteristics of radar sensors.

Keywords: radar sensor, short-range radar system, test equipment, anechoic chamber, delay line.

Иконников В.Н.¹, Канаков В.А.², Пархачёв В.В.²**МЕТОД РЕКОНСТРУКЦИИ ПРОФИЛЯ ТЕМПЕРАТУРЫ
БЫСТРОПРОТЕКАЮЩИХ ПРОЦЕССОВ ПО РЕЗУЛЬТАТАМ
МНОГОКАНАЛЬНЫХ РАДИОМЕТРИЧЕСКИХ ИЗМЕРЕНИЙ**¹*Филиал РФЯЦ-ВНИИЭФ «НИИИС им. Ю.Е. Седакова»,*²*ННГУ им. Н.И. Лобачевского*

В докладе описан метод реконструкции двумерного профиля температуры фронта быстропротекающего газодинамического процесса посредством сплайн-аппроксимации результатов измерения многоканальным радиометром с весовыми коэффициентами, определяемыми диаграммами направленности антенн. Приводятся результаты численного моделирования данного метода для различных модельных распределений яркостной температуры.

Ключевые слова: многоканальный радиометр, яркостная температура, диаграмма направленности, сплайн-аппроксимация.

Для исследования быстропротекающих газодинамических процессов в последние годы активно применяются и развиваются методы микроволнового зондирования: интерферометрия и радиометрия. При помощи одноканальных радиоинтерферометров, разработанных в НИИИС им. Ю.Е. Седакова, решается задача измерения параметров движения фронта детонации вдоль линии визирования, при этом на основе радиоинтерферометра может быть реализован радиометр для измерения температуры исследуемого объекта [1]. Наряду с регистрацией одномерных параметров, актуальны задачи исследования динамики изменения формы объекта и распределения температуры по его поверхности при газодинамических процессах. Для восстановления двумерного поля перемещений используется метод многоканальной радиоинтерферометрии [2]. Однако методы восстановления закона распределения температуры по поверхности газодинамического процесса на основе измерительных данных многоканального радиометра в настоящее время недостаточно разработаны.

По энергетическим соображениям антенная система радиометра миллиметрового диапазона длин волн должна быть расположена вблизи от объекта, а перемещение фронта детонации в ходе эксперимента может составлять несколько десятков длин волн. Исследуемый объект при этом должен оставаться в зоне покрытия диаграмм направленности (ДН) антенн, вместе с тем, ДН не должны быть шире объекта, чтобы не захватывать излучение фона. Таким образом, при приближении фронта детонации к антеннам, размеры и перекрытие областей поверхности ДН будут сильно меняться. Яркостная температура будет усредняться в большом пятне с переменным радиусом, что существенно снизит точность измерений. Двумерный профиль температуры может быть реконструирован по результатам многоканального зондирования с ограниченным количеством каналов путем интерполяции измерительных данных. Для сохранения точности при интерполяции необходимо учитывать эффект изменения углового размера области измерения.

Пусть поверхность с неизвестным распределением температуры движется в направлении антенной системы многоканального радиометра. Форма поверхности и её расположение относительно радиометра должны быть известны в каждый момент времени, например, измерены с помощью многоканального радиоинтерферометра. На рисунке 1а условно показано взаимное расположение антенн радиометра в виде конических рупоров и поверхности, а заштрихованными кругами показаны области,

стягиваемые ДН антенн. Перекрытие ДН и их ширина могут быть произвольными, но не должны выходить за границы объекта исследования.

Пусть $\varphi(x, y)$ – функция, описывающая истинное распределение яркостной температуры по поверхности в некоторой выбранной системе координат. Поскольку яркость излучения может зависеть от угла между нормалью к поверхности и направлением визирования, обозначим эту зависимость $L(\gamma)$, где, γ – угол между нормалью к поверхности в текущей точке (x, y) и направлением на фазовый центр антенны. Обозначим ДН антенн радиометра по мощности $\Psi_i^*(\alpha, \beta)$, где α и β – угловые координаты в некоторой привязанной к антенне системе координат, i – номер антенны.

Тогда яркостная температура f_i , измеренная i -й антенной, может быть выражена через интеграл

$$f_i = K_i \iint_{S_i} \varphi(x, y) L(\gamma) \Psi_i^*(\alpha, \beta) J(x, y, \gamma, \vec{\rho}_i) dx dy, \quad (1)$$

где интегрирование ведётся по пятну поверхности S_i , охваченному ДН, α и β – угловые координаты точки (x, y) в системе координат, связанной с антенной, $J(x, y, \gamma, \vec{\rho}_i)$ – угловой размер, который имеет элементарная площадка интегрирования при наблюдении из фазового центра антенны (якобиан), $\vec{\rho}_i$ – радиус-вектор, проведённый из фазового центра антенны к точке (x, y) . K_i – калибровочный множитель. Введённые обозначения проиллюстрированы на рисунке 1б.

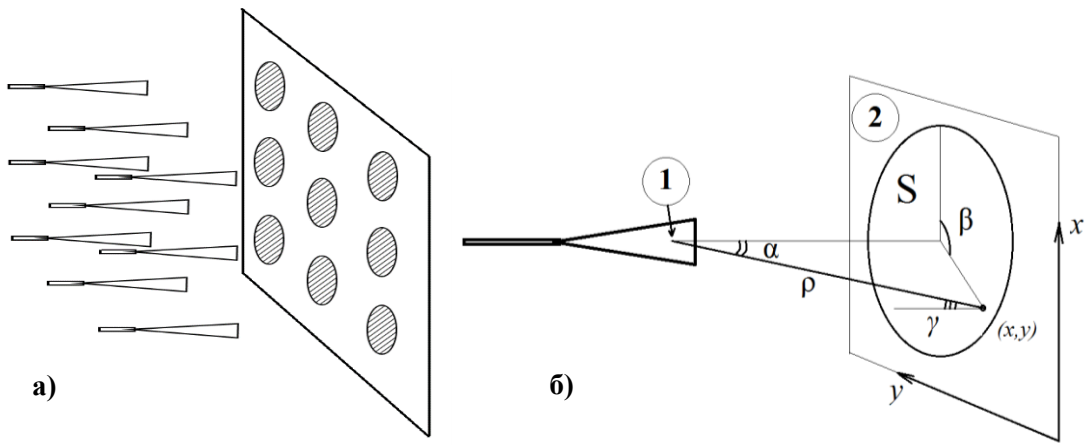


Рис. 1. Взаимное расположение антенн радиометра и объекта исследования (а) и иллюстрация обозначений (б): 1 - фазовый центр антенны, 2 – исследуемый объект

При заданных взаимных положениях поверхности и антенны радиометра множитель $K_i L(\gamma) \Psi_i^*(\alpha, \beta) J(x, y, \gamma, \vec{\rho}_i)$ в интеграле (1) зависит только от x и y . Обозначив его $\Psi_i(x, y)$, интеграл (1) можно переписать в виде:

$$f_i = \iint_{S_i} \varphi(x, y) \Psi_i(x, y) dx dy. \quad (2)$$

Далее будем считать, что функции $\Psi_i(x, y)$ априорно известны.

В случае, когда все ДН антенн радиометра настолько узки, что охватывают области исследуемого объекта S_i , в каждой из которых яркостная температура $\varphi(x, y)$ постоянна, можно считать, что в результате измерения становятся известны точные значения яркостной температуры f_i в определенных точках поверхности. Этих точек столько же, сколько антенн радиометра. Для восстановления приближённого распределения яркостной температуры по поверхности можно использовать процедуру двумерной сплайн-интерполяции или сплайн-аппроксимации [3].

Пусть $f_i = \varphi(x_i, y_i)$, где φ – функция, известная в N точках с координатами (x_i, y_i) , а $i \in [1, N]$. Тогда непрерывная функция $\varphi^*(x, y)$, задаваемая выражением

$$\varphi^*(x, y) = \sum_{i=1}^N C_i r_i^2 \ln r_i^2 + C_{N+1} + C_{N+2}x + C_{N+3}y, \quad (3)$$

$$r_i^2 = (x - x_i)^2 + (y - y_i)^2$$

будет являться сплайн-интерполяцией точек f_i , если коэффициенты C_i удовлетворяют системе уравнений

$$\begin{pmatrix} 0 & \rho_{12} & \dots & \rho_{1N} & 1 & x_1 & y_1 \\ \rho_{21} & 0 & \dots & \rho_{2N} & 1 & x_2 & y_2 \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots \\ \rho_{N1} & \rho_{N2} & \dots & 0 & 1 & x_N & y_N \\ 1 & 1 & \dots & 1 & 0 & 0 & 0 \\ x_1 & x_2 & \dots & x_N & 0 & 0 & 0 \\ y_1 & y_2 & \dots & y_N & 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} C_1 \\ C_2 \\ \dots \\ C_N \\ C_{N+1} \\ C_{N+2} \\ C_{N+3} \end{pmatrix} = \begin{pmatrix} f_1 \\ f_2 \\ \dots \\ f_N \\ 0 \\ 0 \\ 0 \end{pmatrix}, \text{ где:} \quad (4)$$

$$\rho_{ij} = r_{ij}^2 \ln r_{ij}^2$$

$$r_{ij}^2 = (x_i - x_j)^2 + (y_i - y_j)^2$$

В случае, когда яркостная температура внутри области S_i не одинакова, точные значения яркостной температуры в определенных точках поверхности неизвестны. Рассмотрим функцию вида (3). Предположим, что функция $\varphi^*(x, y)$ может точно описать реальное распределение яркостной температуры $\varphi(x, y)$ при некоторых значениях параметров C_i . Подставим в (2) $\varphi^*(x, y)$ вместо $\varphi(x, y)$.

$$f_i^* = \iint_{S_i} \varphi^*(x, y) \Psi_i(x, y) dx dy, \quad (5)$$

где f_i^* – результат измерения. Подставив в (5) выражение для $\varphi^*(x, y)$ из (3), получим систему из N уравнений для коэффициентов C_i :

$$\begin{pmatrix} \rho_{11} & \rho_{12} & \dots & \rho_{1N} & a_1 & b_1 & d_1 \\ \rho_{21} & \rho_{22} & \dots & \rho_{2N} & a_2 & b_2 & d_2 \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots \\ \rho_{N1} & \rho_{N2} & \dots & \rho_{NN} & a_N & b_N & d_N \\ 1 & 1 & \dots & 1 & 0 & 0 & 0 \\ x_1 & x_2 & \dots & x_N & 0 & 0 & 0 \\ y_1 & y_2 & \dots & y_N & 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} C_1 \\ C_2 \\ \dots \\ C_N \\ C_{N+1} \\ C_{N+2} \\ C_{N+3} \end{pmatrix} = \begin{pmatrix} f_1^* \\ f_2^* \\ \dots \\ f_N^* \\ 0 \\ 0 \\ 0 \end{pmatrix}, \text{ где:}$$

$$\rho_{ij} = \iint_{S_i} \Psi_i(x, y) r_j^2 \ln r_j^2 dx dy, \quad (6)$$

$$a_i = \iint_{S_i} \Psi_i(x, y) dx dy,$$

$$b_i = \iint_{S_i} x \Psi_i(x, y) dx dy,$$

$$d_i = \iint_{S_i} y \Psi_i(x, y) dx dy.$$

□

Поскольку в системе уравнений (6) первые N элементов главной диагонали не равны нулю, функция, задаваемая уравнением (3), будет являться сплайн-аппроксимацией исходных точек f_i и позволяет восстановить распределение яркостной температуры по измеренным значениям.

Для оценки величины ошибки, которая возникает при восстановлении распределения яркостной температуры, рассмотрим три случая, когда истинное распределение яркостной температуры $\varphi(x, y)$ описывается частью наклонной плоскости, частью эллиптического параболоида и частью гиперболического параболоида (рисунок 2). Величину ошибки определим с помощью численного моделирования. Будем считать, что

исследуемый объект представляет собой часть плоскости размером 1×1 условных единиц длины. Введём на нём прямоугольную систему координат XU : $x \in [0,1]$, $y \in [0,1]$. Пусть радиометр имеет девять каналов, оси симметрии ДН перпендикулярны плоскости и пересекаются с ней в точках с координатами (x_i, y_i) : $(0.25, 0.25)$, $(0.25, 0.5)$, $(0.25, 0.75)$, $(0.5, 0.25)$, $(0.5, 0.5)$, $(0.5, 0.75)$, $(0.75, 0.25)$, $(0.75, 0.5)$, $(0.75, 0.75)$.

Сначала определим функцию $\Psi_i(x, y)$ как константу $1/(\pi R^2)$ в круге с центром (x_i, y_i) и радиусом R . Найдём величину ошибки для R равных 0.0001 , 0.1 и 0.25 . В первом случае радиус меньше шага численного интегрирования, который был равен 0.001 , что позволяет промоделировать случай бесконечно узких ДН. Во втором случае ДН от разных антенн не пересекаются, а в третьем настолько широки, что пересекаются.

Моделирование заключается в вычислении девяти значений яркостной температуры на каждом из заданных распределений $\varphi(x, y)$ по формуле (2), которые соответствуют измеренным радиометром величинам. Далее, из системы (6) найдём коэффициенты сплайн-поверхности $\varphi^*(x, y)$. Ошибкой будем считать усредненное по площади среднеквадратичное отклонение истинного распределения от найденной сплайн-поверхности, нормированное на разброс яркостной температуры:

$$E = \frac{\sqrt{\frac{1}{S} \iint_S [\varphi(x, y) - \varphi^*(x, y)]^2 dx dy}}{\max \varphi(x, y) - \min \varphi(x, y)} \quad (7)$$

В табл. 1 приведены результаты численного моделирования. Представлены значения ошибки E для трёх типов распределения $\varphi(x, y)$ и трёх значений радиуса стягиваемой ДН области поверхности R .

Графики исходных (красный) и восстановленных (9 измерительных каналов, с учётом ДН – синий) распределений температуры при $R = 0.25$ приведены на рис. 2. Точками на графиках показаны расположение антенн радиометра.

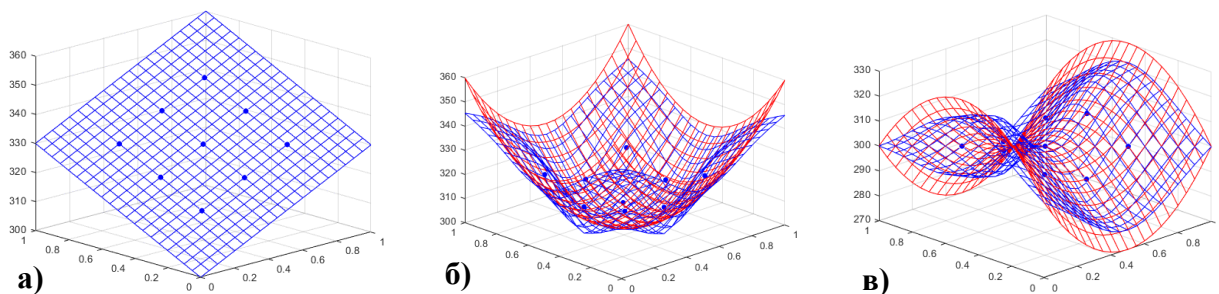


Рис. 2. Исходные (красный) и восстановленные (синий) распределения температуры: плоскость (а), эллиптический параболоид (б), гиперболический параболоид (в)

Таблица 1

Ошибки восстановления распределения яркостной температуры с учётом ДН, 9 каналов

	$R = 0.0001$	$R = 0.1$	$R = 0.25$
Плоскость	0	0	0
Эллиптический параболоид	0.26	0.22	0.12
Гиперболический параболоид	0.25	0.22	0.14

Из табл. 1 видно, что наклонная плоскость описывается сплайн-поверхностью точно (см. рис. 2а), вне зависимости от величины радиуса R . Кроме того, средняя ошибка восстановления значения яркостной температуры в каждой точке много меньше разброса яркостной температуры по всей поверхности, и она убывает с увеличением ширины ДН.

Для сравнения рассмотрим случай, когда измеренные значения яркостной температуры подставляются в систему уравнений (4) для определения коэффициентов сплайн-поверхности без учёта ДН. Получившиеся ошибки представлены в табл. 2. Из сравнения табл. 1 и 2 видно, что, за исключением случая наклонной плоскости, ошибка аппроксимации без учёта ДН выше, чем с их учётом. Таким образом было достигнуто повышение точности сплайн-аппроксимации профиля температуры при замене узлов интерполяции областями интегрирования с конечными размерами и известным весовым распределением.

Таблица 2

Ошибки восстановления распределения яркостной температуры без учёта ДН, 9 каналов

	$R = 0.0001$	$R = 0.1$	$R = 0.25$
Плоскость	0	0	0
Эллиптический параболоид	0.26	0.25	0.22
Гиперболический параболоид	0.25	0.25	0.25

Также рассмотрим, как ошибка при увеличении числа измерительных каналов. Количество антенн увеличим до 16 и расположим их в узлах квадратной сетки в точках (0.2, 0.2), (0.2, 0.4), (0.2, 0.6), (0.2, 0.8), (0.4, 0.2), (0.4, 0.4), (0.4, 0.6), (0.4, 0.8), (0.6, 0.2), (0.6, 0.4), (0.6, 0.6), (0.6, 0.8), (0.8, 0.2), (0.8, 0.4), (0.8, 0.6), (0.8, 0.8). Как и ранее, положим R равным 0.0001, 0.1 и 0.25. Коэффициенты сплайн-поверхности рассчитываются по системе уравнений (4). Результаты представлены в табл. 3.

Таблица 3

Ошибки восстановления распределения яркостной температуры без учёта ДН(16 каналов)

	$R=0.0001$	$R=0.1$	$R=0.25$
Плоскость	0	0	0
Эллиптический параболоид	0.16	0.13	0.08
Гиперболический параболоид	0.17	0.15	0.11

Сравнив табл. 1 и 3, отметим, что увеличение числа измерительных каналов привело к значительному уменьшению ошибки, что объясняется наличием большей апостериорной информации при определении формы сплайн-поверхности. При этом величина ошибки восстановления в случае 9 каналов с учетом ДН ($R = 0.25$) становится сопоставимой со случаем 16 каналов без учета ДН ($R=0.1$).

В работе [4] обсуждается применение в многоканальной радиоинтерферометрии излучателей, формирующих гауссову осесимметричную ДН. Рассмотрим работоспособность метода в случае, если функция $\Psi_i(x, y)$ имеет вид двумерного гауссовского распределения в круге с центром (x_i, y_i) и радиусом R , равным 2σ :

$$\Psi_i(x, y) = \frac{1}{\pi R^2/2} \exp\left(-\frac{(x-x_i)^2+(y-y_i)^2}{R^2/2}\right). \quad (8)$$

В табл. 4 приведены результаты моделирования для 9 измерительных каналов, R выберем равным 0.1 и 0.25.

Ошибки восстановления распределения яркостной температуры с учётом ДН (гауссовское распределение), 9 каналов

	$R = 0.1$	$R = 0.25$
Плоскость	0	0
Эллиптический параболоид	0.24	0.15
Гиперболический параболоид	0.24	0.18

Из сравнения табл. 1 и 4, видно, что ошибки восстановления возрастают по сравнению со случаем равномерного распределения, однако убывают с увеличением ширины ДН. Таким образом, предложенный метод сохраняет свою работоспособность в случае гауссовской функции $\Psi_i(x, y)$.

Исходя из результатов численного моделирования можно заключить, что описанный метод позволяет восстанавливать двумерные распределения яркостной температуры, измеряемые с помощью многоканальных радиометров, посредством сплайн-аппроксимации с точностью, превосходящей точность обычной сплайн-интерполяции, проводимой без учёта влияния ДН антенн.

Библиографический список

1. Канаков В.А. Состояние и перспективы развития микроволновой радиоинтерферометрии для диагностики газодинамических процессов/ В.А. Канаков, С.В. Катин, Н.С. Корнев, А.Л. Михайлов, А.В. Назаров, Ю.И. Орехов, А.В. Родионов, В.Н. Хворостин // Антенны. – 2016. – №1 (221). – С. 49-54.
2. Катин С.В. Многоканальная радиоинтерферометрия – метод диагностики изменения фронтов ударноволновых и детонационных процессов. Концепция и экспериментальное подтверждение/ С.В. Катин, Н.И. Дмитриев, Ю.И. Орехов, А.Л. Михайлов, В.Ф. Взятыхшев, В.А. Канаков, А.В. Родионов, В.Н. Хворостин// Труды Международной конференции «Экстремальные состояния вещества. Детонация. Ударные волны». – Саров: РФЯЦ-ВНИИЭФ. – 2009. – С. 617–621.
3. Де Бур К., Практическое руководство по сплайнам: Пер. с англ. – М.: Радио и связь, 1985. – 304 с.
4. Катин С.В. Стратегия выбора структуры и параметров волновой подсистемы многоканальных радиоинтерферометров/ С.В. Катин, В.Ф. Взятыхшев, Ю.И. Орехов, В.Н. Хворостин // Труды Международной конференции «Экстремальные состояния вещества. Детонация. Ударные волны». – Саров: РФЯЦ-ВНИИЭФ. – 2011. – С.670-674.

V. N. Ikonnikov¹, V. A. Kanakov², V. V. Parkhachev²

THE METHOD OF RESTORATION OF THE TEMPERATURE PROFILE OF FAST-FLOWING PROCESSES BY RESULTS OF MULTICHANNEL RADIOMETRIC MEASUREMENT

¹*Branch of the FSUE Russian Federal Nuclear Center – All-Russian Research Institute of Experimental Physics Research Institute of Measuring Systems n.a. Yu.Ye. Sedakov, Nizhny Novgorod, Russia*

²*Lobachevsky State University of Nizhny Novgorod, Russia*

The article describes the method for restoration of the two-dimensional temperature profile of the front of a fast-flowing gas-dynamic process by spline approximation of results of multichannel radiometric measurement with weight coefficients determined by antenna radiation patterns. The results of numerical simulation of this method for various model distributions of the brightness temperature are presented.

Keywords: multichannel radiometer, radiation pattern, brightness temperature, spline approximation.

УЛУЧШЕНИЕ ЧАСТОТНОГО РАЗРЕШЕНИЯ МЕТОДА MULTITAPER SPECTRUM ESTIMATION НА ОСНОВЕ ПОЛИГАРМОНИЧЕСКОЙ ЭКСТРАПОЛЯЦИИ

Нижегородский государственный университет им. Н.И. Лобачевского

Возможны различные подходы к построению спектрограмм сигнальных процессов на основе применения разных методов оценивания спектральной плотности мощности сигналов. В качестве основного в данной работе рассмотрен метод “Multitaper Spectrum Estimation”, позволяющий снизить ограничения классических периодограммных и коррелограммных методов спектрального оценивания, такие как ограничение частотного разрешения и «утечка» энергии из главного лепестка спектральной линии в боковые лепестки в частотной области. Исследуется возможность применения полигармонической экстраполяции с целью улучшения качества получаемых спектрограмм сигналов.

Ключевые слова: спектральная плотность мощности, спектрограмма, преобразование Фурье, Multitaper spectrum estimation, полигармоническая экстраполяция.

Введение

Спектрограммы занимают важное место в современной цифровой обработке сигналов. Спектрограммы визуализируют изменение спектра из частот сигнала во времени и получили широкое применение в различных научных и технических приложениях, в частности, в областях обработки звуков, музыки, речи, в радио- и гидролокации, сейсмологии [1, 2]. При построении спектрограмм обычно применяется оконное преобразование Фурье: спектр вычисляется от последовательных окон сигнала, и каждый из этих спектров образует столбец в спектрограмме.

Такой метод спектрального анализа эффективен в вычислительном плане и подходит для большого количества различных сигнальных процессов, однако имеет ряд недостатков, которые наиболее сильно проявляются при обработке относительно коротких выборок сигналов, с которыми чаще всего сталкиваются на практике. В результате надежность оценок спектрограмм, основанных на простом использовании алгоритмов дискретного преобразования Фурье, значительно понижается из-за дисперсии оценки спектральной плотности мощности (СПМ) анализируемых сигналов.

Нелинейные алгоритмы спектрального оценивания [1, 3-5] позволяют существенно снизить влияние боковых лепестков, повысить спектральное разрешение при обработке коротких выборок сигналов, однако им также присущи определенные недостатки, в частности, требуется выбор адекватного порядка модели, они менее устойчивы к влиянию шума.

МЕТОД ОЦЕНКИ СПМ «MULTITAPER SPECTRUM ESTIMATION»

В работе для построения спектрограмм рассмотрен метод Multitaper Spectrum Estimation, позволяющий повысить достоверность оценок СПМ и ослабить ограничения традиционных линейных методов, основанных на применении алгоритма быстрого преобразования Фурье. Multitaper Spectrum Estimation также относится к линейным методам, при этом он позволяет улучшить оценку СПМ, устраняя утечки энергии в боковые лепестки, и уменьшить дисперсию спектральной оценки [6, 7]. Метод Multitaper предполагает использование дискретных вытянутых сфероидальных последовательностей (функции Слепяна) в качестве специальных окон, вместо обычного прямоугольного окна [8, 9]. На рис. 1 представлен вид первых 5-и таких окон.

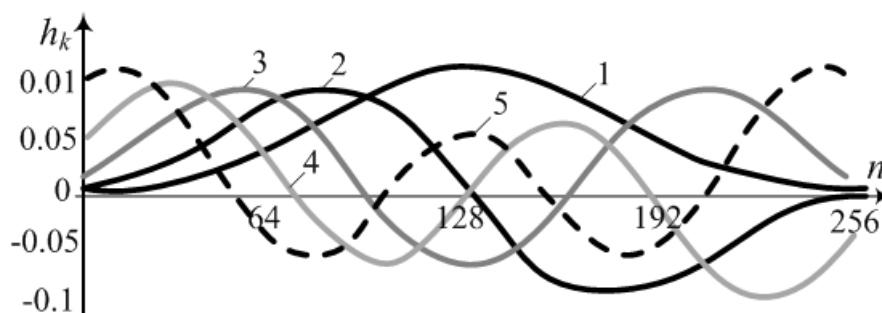


Рис. 1. Дискретные вытянутые сфероидальные последовательности

При построении спектрограммы сигнала для каждой временной последовательности $x[n]$ оценка спектра вычисляется следующим образом

$$S[m] = \left| \sum_{n=0}^{N-1} x[n] h_k[n] e^{-j2\pi km/N} \right|^2, \quad 0 \leq m \leq N-1, \quad (1)$$

Сигнал предварительно умножается на специальное окно $h_k[n]$, затем от этого берётся преобразование Фурье. Спектральная оценка получается путём усреднения нескольких полученных оконных БПФ

$$P_{mtr}[m] = \frac{1}{K} \sum_{n=0}^{N-1} S_k(m). \quad (2)$$

Полигармоническая экстраполяция пространственно-временных рядов на основе спектральных представлений

Для сравнения методов при построении спектрограмм использован сигнал, состоящий из 5 синусоид с различными амплитудами, частотами и начальными фазами в белом гауссовом шуме, отношение сигнал/шум -3 дБ. Пример построения спектральной оценки и, соответственно, спектрограммы модельного сигнала приведен на рис. 3 а. Спектральные линии сигналов, полученные с использованием метода “Multitaper Spectrum Estimation”, довольно широкие, если сравнивать, например, с линиями, получающимися в результате использования периодограммного метода. От толщины спектральных линий напрямую зависят точность и качество получаемых спектрограмм, чем толще линии – тем ниже разрешающая способность метода.

С целью исследования возможности применения полигармонической экстраполяции для повышения спектрального разрешения получаемых спектрограмм был реализован алгоритм, расширяющий реализацию сигналов (длину окна) вперед и назад перед процедурами получения спектра. В работе используется упрощенный алгоритм прогноза на основе спектральных представлений [10]. Алгоритм экстраполяции основан на оценке амплитуд и фаз, гармонических составляющих сигнала в предположении о сохранении периодических закономерностей процесса и медленности изменения спектральных характеристик на интервале прогноза и на интервале предыстории. Аналитическое выражение для последовательности операций алгоритма экстраполяции имеет вид:

$$S_{extr}(t) = M \times F^{-1}[F(S_{base}(t)) \times |F(S_{base}(t))|], \quad (3)$$

где $S_{base}(t)$ – отрезок с отсчетами предыстории, $S_{extr}(t)$ – отрезок предыстории, расширенный на удвоенный интервал экстраполяции; F^{-1} и F – соответственно операторы обратного и прямого Фурье-преобразования; M – масштабирующий множитель.

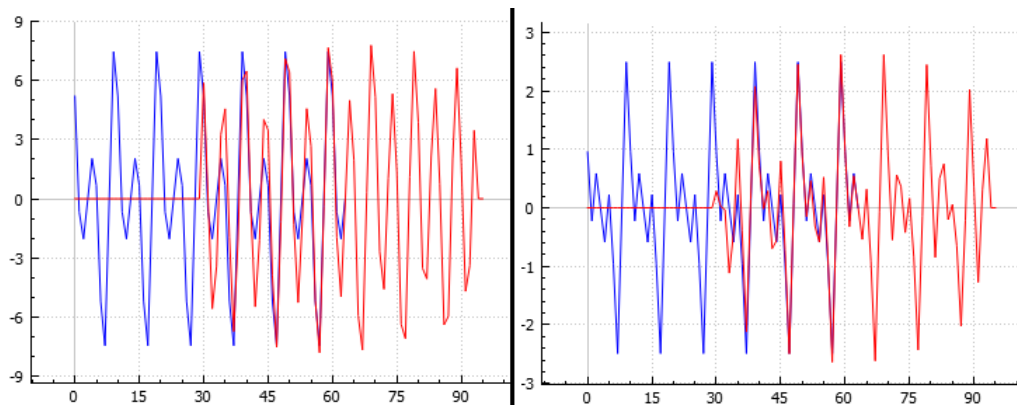


Рис. 2. Результаты полигармонической экстраполяции сигналов вперед

Примеры работы алгоритма экстраполяции вперед приведены на рис. 2. Таким образом, при построении спектрограммы после разбиения сигнала на временные окна для каждого окна применяется экстраполяция сигнала **вперед и назад**, тем самым увеличивая размер этого окна за счет прогнозирования отсчетов сигнала. После этого от последовательных расширенных окон сигнала вычисляется спектральная оценка. Данная процедура может быть использована при вычислении различных спектральных оценок сигнала. На рис. 3 представлены спектральные оценки и спектрограммы модельного сигнала, полученные методом “Multitaper Spectrum Estimation” без использования экстраполяции и с применением полигармонической экстраполяции.

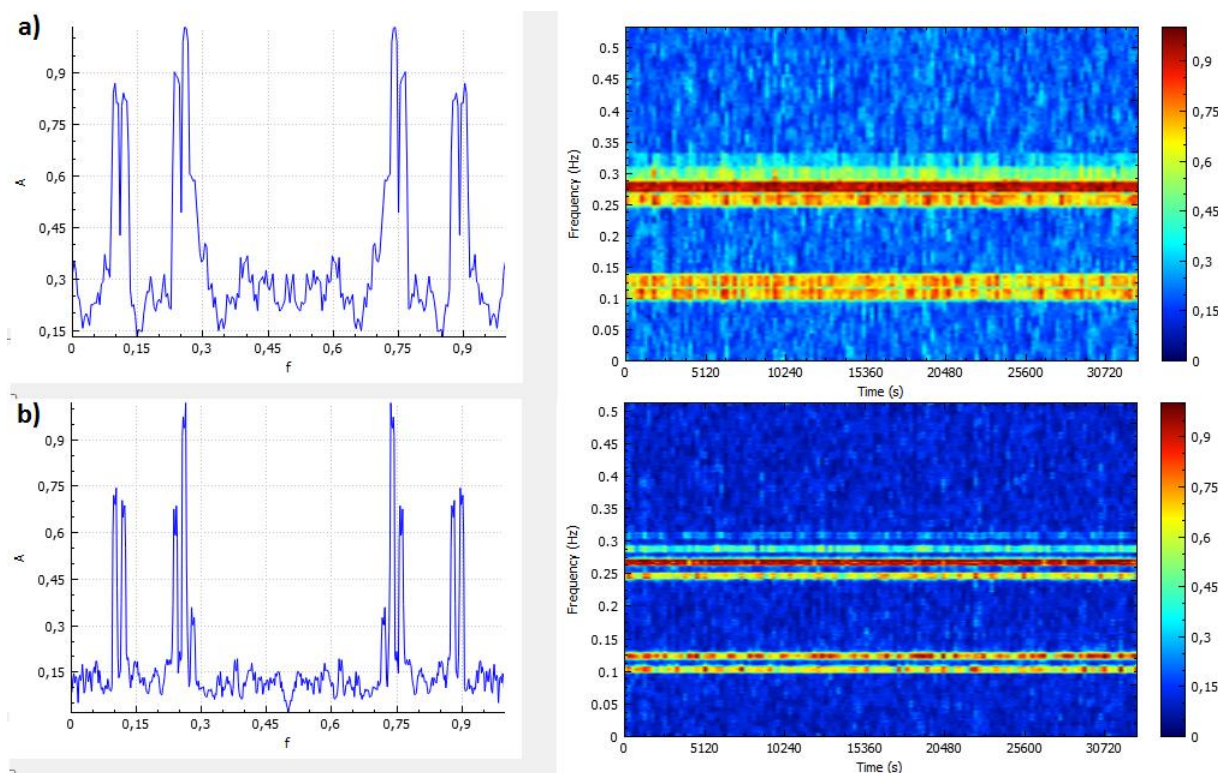


Рис. 3. Спектрограммы, полученные методом “Multitaper Spectrum Estimation”, отношение сигнал/шум –3 дБ (а - без использования экстраполяции, б – с применением полигармонической экстраполяции)

Как показывают результаты моделирования, применение метода полигармонической экстраполяции способствовало повышению спектрального разрешения. Сравнение результатов построения спектрограмм сигналов без использования экстраполяции и с применением полигармонической экстраполяции (рис.

3) показывает, что спектральные линии после применения экстраполяции стали значительно тоньше, существенно уменьшился эффект наложения спектральных линий друг на друга. При этом сохраняется **надежность и достоверность оценок СПМ, построенных методом “Multitaper Spectrum Estimation” с использованием алгоритма полигармонической экстраполяции,** и ослабляются ограничения традиционных линейных методов. Алгоритм так же устойчив к влиянию наложенного на сигнал шума.

Библиографический список

1. Марпл.-мл. С.Л. Цифровой спектральный анализ и его приложения / С.Л. Марпл.-мл. – М.: Мир, 1990. – 584 с.
2. Сергиенко А.Б. Цифровая обработка сигналов / А.Б. Сергиенко – СПб.: БХВ-Петербург, 2011. – 768 с.
3. Сорохтин М.М. Применение нелинейного спектрального оценивания в задаче определения взаимной временной задержки сигналов / М.М. Сорохтин, Е.М. Сорохтин, О.А. Морозов, В.Р. Фидельман // Известия высших учебных заведений. Радиофизика. Т. 50, № 4. 2007. С. 357-363.
4. Морозов О.А. Эффективный вычислительный алгоритм реализации метода максимальной энтропии в задачах обращения свертки / О.А. Морозов, Т.Г. Рыжкова, В.Р. Фидельман // Известия высших учебных заведений. Радиофизика. Т. 45, № 8. 2002. С. 722-729.
5. Клименченко П.В. Авторегрессионный алгоритм Берга для обнаружения целей и определения их скоростей на фоне пассивных помех, основанный на спектральных и статистических различиях целей и помех / П.В. Клименченко, В.Н. Жураковский // Радиостроение, № 4. 2017. С. 1-15.
6. Thomson D.J. Spectrum estimation and harmonic analysis / D.J. Thomson // Processings of the IEEE, V. 70, № 9. 1982.
7. Van Dronghelen W. Multitaper Power Spectrum Estimation / W. Van Dronghelen. 2014. URL:<https://pdfs.semanticscholar.org/752d/1a551b96559458064323eb3de7faaaef4c4e.pdf>.
8. Gruenbacher D.M. A Simple Algorithm for Generating Discrete Prolate Spheroidal Sequences / D.M. Gruenbacher, D.R. Hummels // IEEE Transactions on signal processing. V. 42, № 11. 1994.
9. Лимарев А.Е. Быстрый алгоритм формирования дискретных последовательностей Слепяна / А.Е. Лимарев // Теория и техника радиосвязи, № 2. 2013.
10. Евсеев А.П. Экстраполяция (прогнозирование) пространственно-временных рядов на основе спектральных представлений / А.П. Евсеев, Д.А. Евсеев, В.В. Баданов // Вестник Нижегородского университета им. Н.И. Лобачевского. Серия Радиофизика. №1. 2004. С 249-255.

A.N. Andreev, O.A. Morozov

IMPROVING THE FREQUENCY RESOLUTION OF THE MULTITAPER SPECTRUM ESTIMATION METHOD BASED ON POLYHARMONIC EXTRAPOLATION

There are various approaches to constructing spectrograms of signal processes based on using of different methods for estimating the spectral power density of signals. The method “Multitaper Spectrum Estimation” is considered as the main one in this work, which allows to reduce the limitations of classical periodogram and correlogram methods of spectral estimation, such as limiting the frequency resolution and “leakage” of energy from the main lobe of the spectral line to the side lobes in the frequency domain. The possibility of using polyharmonic extrapolation to improve the quality of spectrograms is investigated.

Keywords: power spectral density, spectrogram, Fourier transform, Multitaper spectrum estimation, polyharmonic extrapolation.

СХЕМОТЕХНИЧЕСКАЯ МОДЕЛЬ АСИММЕТРИЧНЫХ ЦИРКУЛЯТОРОВ НА СОСРЕДОТОЧЕННЫХ ЭЛЕМЕНТАХ

Филиал ФГУП «РФЯЦ-ВНИИЭФ» «ФНПЦ НИИИС им. Ю.Е. Седякова»

Объектом исследования являются ферритовые циркуляторы на сосредоточенных элементах с асимметрией плеч. В работе представлен общий схемотехнический подход к расчету узкополосных и широкополосных асимметричных ферритовых развязывающих устройств на сосредоточенных элементах. Рассматривается два вида асимметрии циркуляторов, обусловленной различием в характеристиках нагрузочных сопротивлений плеч циркулятора и определяемой технологическими особенностями сборки системы переплетения проводников ферритового Y-сочленения. В расчете учитываются потери в феррите. Методика схемотехнического расчета интегрирована в САПР СВЧ устройств AWR Microwave Office.

Ключевые слова: асимметричный циркулятор на сосредоточенных элементах, система переплетённых проводников, паразитные схемные параметры.

Введение

Ферритовые циркуляторы/вентили на сосредоточенных элементах широко применяются в приемопередающей аппаратуре метрового и дециметрового диапазонов для защиты мощных источников радиосигналов от переотражений в СВЧ трактах, для обеспечения их стабильности в случае изменяющейся во времени нагрузки и для развязки передатчика, приемника и антенны.

Основным функциональным элементом в циркуляторах на сосредоточенных элементах является ферритовое Y-сочленение, которое образовано двумя ферритовыми элементами и расположенной между ними системой переплетённых проводников (рис.1). Система переплетения образована шестью проводниками: по два на каждое плечо циркулятора. Сдвоенные проводники расположены под углом в 120 градусов по отношению к соседним. Проводники в точках пересечения изолированы друг от друга. Переплетение проводников обеспечивает равномерное возбуждение магнитного поля в нижнем и верхнем ферритовых элементах. Проводники каждого порта электрически объединены на концах и подключаются по выходам к общей точке Y-сочленения (обратная металлизация ферритовых элементов). В каждое плечо циркулятора вводятся согласующие цепи. Для расширения полосы рабочих частот добавляют цепи общего контура: между общей точкой ферритового Y-сочленения и заземлением устройства.

Традиционно, расчет таких устройств основан на рассмотрении симметричной схемы замещения ферритового Y-сочленения и базируется на методах линейной алгебры, применяемых к матрице импедансов циркулятора (метод собственных возбуждений) [1]. Физическое обоснование этого подхода основано на факте существования в ферритовых материалах собственных волн с правым и левым направлениями вращения плоскости поляризации.

При феноменологическом расчете циркуляторов на сосредоточенных элементах принимают ряд стандартных физических допущений [2]. Так, предполагается, что в случае возбуждения только одного порта циркулятора высокочастотное магнитное поле однородно распределено по всей поверхности ферритовых дисков и направлено перпендикулярно проводникам. Электрическая длина проводников столь мала, что вся высокочастотная энергия заключается в магнитном поле. Потери в феррите отсутствуют, и амплитуды собственных волн с круговой поляризацией одинаковы. Коэффициент магнитной связи индуктивностей по высокочастотному магнитному потоку равен единице.

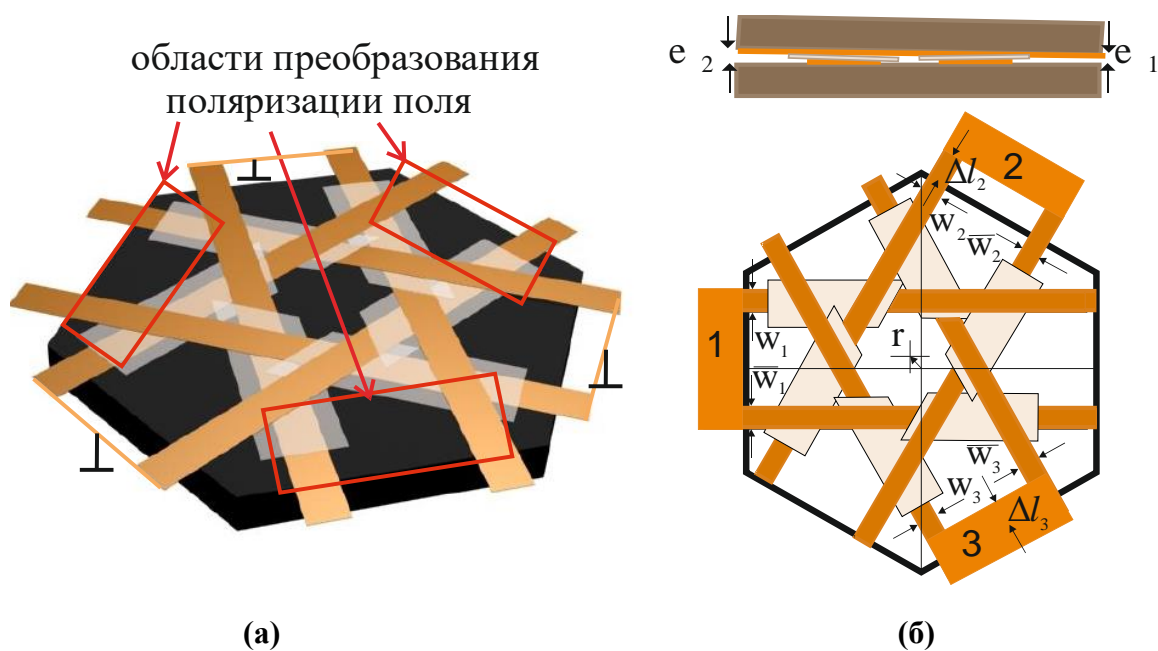


Рис. 1. Ферритовое Y-сочленение (верхний ферритовый элемент не показан) (а), схема конструкторско-технологических погрешностей изготовления системы переплетенных проводников (б)

В рамках принятых физических допущений для собственных типов возбуждения ферритового Y-сочленения с вращением фазы сигналов на портах на 120 градусов при обходе портов по/против часовой стрелки во всем объеме ферритовых элементов возбуждаются волны с круговой поляризацией электромагнитного поля. Для синфазного собственного возбуждения в центре ферритовых элементов реализуется режим короткого замыкания.

При расчете ферритовых циркуляторов, основанном на феноменологической модели Y-сочленения, паразитные элементы конструкции не учитываются. Анализ основывается на эвристических и эмпирических эквивалентных схемах. Поэтому такие исследования носят характер оценки предельно достижимых характеристик циркуляторов. Вместе с тем паразитные схемные параметры существенно влияют на реально достижимые рабочие характеристики и требуют учета, особенно при проектировании широкополосных и сложно функциональных устройств.

Условия применения циркуляторов в составе приемопередающих модулей часто сопряжены с асимметрией параметров нагрузок плеч устройства. А технологические погрешности в изготовлении топологических элементов ферритового Y-сочленения приводят к дополнительной асимметрии. Такая внешняя и внутренняя асимметрии влекут за собой ухудшение рабочих характеристик устройства. Для компенсации асимметрии на практике вводят подстроечные элементы.

Вместе с тем, в технике ферритовых развязывающих устройств известны примеры технических решений с ярко выраженной асимметрией структуры системы переплетения проводников [3, 4]. Расположение проводников соседних портов ферритового вентиля под углом, отличным от 120 градусов при определенных условиях приводит к снижению вносимых потерь и расширению полосы рабочих частот. Снижение потерь достигается за счет уменьшения амплитуды, возбуждаемой в феррите собственной волны с правым направлением вращения плоскости поляризации, для которой на резонансной частоте имеет место резонанс поглощения. При этом терминальное сопротивление третьего порта циркулятора отличается от стандартного значения 50 Ом. Теоретической модели, позволяющей простыми методами рассчитывать такие асимметричные устройства на данный момент не предложено. Поэтому анализ асимметричных циркуляторов является актуальной и перспективной задачей.

В данной работе предложена схмотехническая модель приближенного расчета циркуляторов на сосредоточенных элементах с произвольной асимметрией. Внутренняя асимметрия задается через паразитные реактивные элементы эквивалентной схемы циркулятора. Внешняя асимметрия задается различием в значениях терминальных нагрузок плеч циркулятора. В статье представлены результаты решения задачи компенсации асимметрии частотных характеристик плеч широкополосного циркулятора за счет вариации номиналов согласующих LC-цепей по критерию соответствия заданному уровню рабочих характеристик. Проводится расчет компенсирующей цепи асимметричной схемы вентиля с емкостью в третьем плече, образуемой мощным нагрузочным резистором.

Приближенный анализ влияния внутренней асимметрии ферритового Y-сочленения

Основные производственные причины внутренней асимметрии сборки системы переплетения проводников интуитивно понятны (рис. 1б). К ним относятся погрешности изготовления ширины Δw_1 - Δw_3 и длины проводников Δl_1 - Δl_3 , погрешность центрирования сборки (r) и магнита, а также его неплоскостность (Δe_{12}) по отношению к ферритам, неравномерность прижатия ферритовых дисков, разные толщины прокладок.

Эти конструкторско-технологические погрешности приводят к разбросу номиналов элементов эквивалентной электрической схемы рис. 2а [5]. В классическую феноменологическую схему замещения ферритового Y-сочленения (рис. 2б в отсутствие потерь) добавлены полуэмпирические паразитные элементы. При этом собственная индуктивность магнитносвязанных проводников L_0 определяет собственные индуктивности идеальной системы переплетения при вращающихся собственных возбуждениях, как $L_{P/M} = \frac{3}{2} L_0 \operatorname{Re} \mu_{\pm}$, где μ_{\pm} - комплексные эффективные магнитные проницаемости волн с правым и левым направлениями вращения плоскости поляризации. Мнимые части μ_{\pm} задают эквивалентные собственные сопротивления потерь в феррите

$R_{P/m}$. Сопротивление R_o определяет активные потери в металлических проводниках, в общем случае, с учетом скин-эффекта. В схеме учитываются паразитные емкости связи проводников в точках их пересечения C_{C12} , C_{C23} , C_{C13} . Паразитные индуктивности нецентральных областей проводников условно разделены на две части: соответствующие периферии ферритового элемента с квази-линейной поляризацией электромагнитного поля L_{iL} , и соответствующие областям с эллиптической поляризацией электромагнитного поля L_{iE} , $i=1 \dots 3$.

На качественном уровне можно утверждать, что неравномерность прижатия ферритовых дисков к системе переплетения приводит к асимметрии паразитной емкости связи между проводниками, разделенными фторопластовыми прокладками, а радиальная неравномерность поля подмагничивания, на фоне неплоскостности магнитов относительно ферритовых дисков, приводит к асимметрии паразитной индуктивности в периферийной области феррита. Индуктивность L_0 определяется, как индуктивность системы двух связанных симметрично запитанных полосковых линий длиной, практически равной диаметру ферритового диска. Индуктивности L_{iE} и L_{iL} определяются через эффективную магнитную проницаемость $\operatorname{Re}(\mu_{\text{eff}})$ волны, распространяющейся перпендикулярно направлению намагничивания. Численно они рассчитываются в зависимости от длины начального участка проводников над ферритовым элементом до первого перекрестия проводников. Сопротивления потерь в феррите определяются из соотношений: $R_{iL/iE} = \omega L_{iL/iE0} |\operatorname{Im}(\mu_{\text{eff}})|$,

$R_{P/M} = \omega \frac{3}{2} L_0 |\operatorname{Im} \mu_{\pm}|$. Область преобразования линейной поляризации падающей волны в волну круговой поляризации, на которой происходит функционирование циркулятора, обозначена на рис. 1 условно.

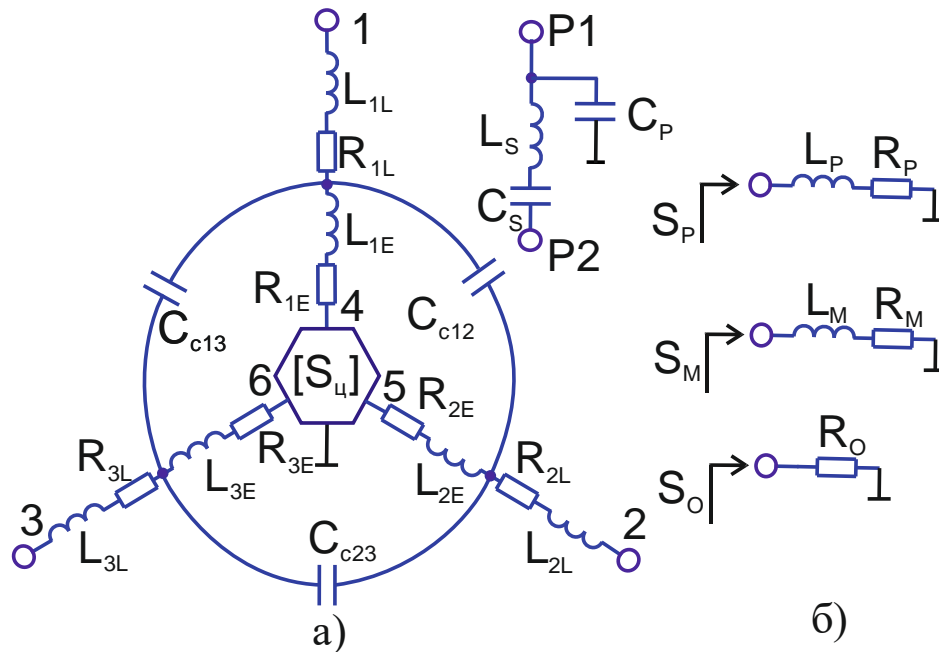


Рис. 2. Эквивалентная схема шестиполюсника паразитных параметров системы переплетения проводников (а) и собственные цепи функциональной части переплетения с матрицей рассеяния $[S_c]$ (б)

Приближенный анализ влияния асимметрии точки установки общего контура

В широкополосных циркуляторах применяется общий контур из последовательной LC цепи (рис. 3а). Асимметрию точки пайки индуктивности к ферритам и центрирования установки емкости на общую землю устройства предложено учитывать геометрическим способом: считать разницу длин путей протекания токов с асимметрией и без неё. Тогда матрица импедансов общего контура [5] может быть представлена в простом виде:

$$Z_g = Z_c \left(\begin{bmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{bmatrix} + K_{as} \begin{bmatrix} \Delta g_{11} & \Delta g_{12} & \Delta g_{13} \\ \Delta g_{21} & \Delta g_{22} & \Delta g_{23} \\ \Delta g_{31} & \Delta g_{32} & \Delta g_{33} \end{bmatrix} \right), \quad (1)$$

где $\Delta g_{11} = \sqrt{(x-1)^2 + y^2} - 1$, $\Delta g_{12} = \Delta g_{21} = \sqrt{(x+0.5)^2 + (y + \sqrt{3}/2)^2} + \sqrt{(x-1)^2 + y^2} - 2$,

$\Delta g_{13} = \Delta g_{31} = \sqrt{(x+0.5)^2 + (y - \sqrt{3}/2)^2} + \sqrt{(x-1)^2 + y^2} - 2$,

$\Delta g_{23} = \Delta g_{32} = \sqrt{(x+0.5)^2 + (y - \sqrt{3}/2)^2} + \sqrt{(x+0.5)^2 + (y + \sqrt{3}/2)^2} - 2$,

Z_c – импеданс последовательного резонансного контура.

Коэффициент асимметрии K_{as} можно называть удельным (на единицу длины) высокочастотным сопротивлением обратной стороны ферритовых элементов. Эта величина имеет смысл лишь в верхней части дециметрового диапазона длин волн.

Схемная модель асимметричного циркулятора

В широкополосных циркуляторах используются полоснорасширяющие согласующие цепи [6] (P1-P2 на рис.1а), устанавливаемые в каждое плечо циркулятора, и цепи общего контура (рис.3а). S-матрица рассеяния получающегося шестиполюсника может быть представлена в блочно-матричном виде (рис.1а). Общая схемная модель асимметричного циркулятора показана на рис. 3б. Результирующие характеристики асимметричного циркулятора получаются в результате матричных операций.

$$S_R = S_n + C_0 S_C (E - S_O S_C)^{-1} C_n. \quad (2)$$

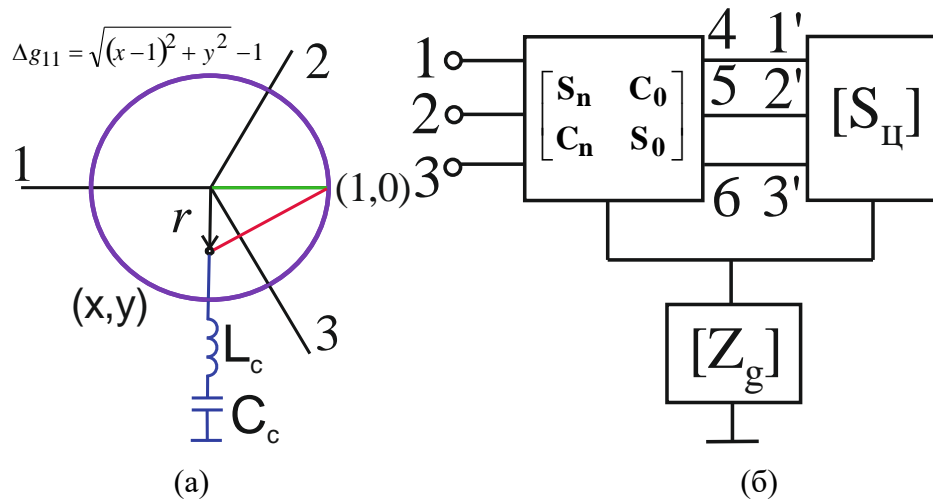


Рис. 3. Схема учета асимметрии точки установки цепи общего контура на обратную сторону ферритового диска (а) и общая схемная модель асимметричного циркулятора (б)

По матрице $[S_R]$ рассчитывается матрица импедансов $Z_R = (E + S_R)^{-1}(E - S_R)$. Матрица импедансов циркулятора с асимметричным общим контуром определяется по формуле последовательного соединения схем: $Z = Z_g + Z_R$. От неё переходим к итоговой матрице рассеяния: $S = (E + Z)^{-1}(E - Z)$.

Матрица рассеяния циркулятора с нагрузкой в третьем плече, для которой имеет место отличие коэффициента отражения (Γ) от нуля, определяется согласно выражениям:

$$S_{11}^V = S_{11} + S_{13}(1 - \Gamma S_{33})^{-1} \Gamma S_{31}; S_{12}^V = S_{12} + S_{13}(1 - \Gamma S_{33})^{-1} \Gamma S_{32}; \quad (3)$$

$$S_{21}^V = S_{21} + S_{23}(1 - \Gamma S_{33})^{-1} \Gamma S_{31}; S_{22}^V = S_{22} + S_{23}(1 - \Gamma S_{33})^{-1} \Gamma S_{32}.$$

Результаты тестирования схемотехнической модели широкополосного асимметричного циркулятора в САПР AWR Microwave Office

Методика расчета асимметричного циркулятора, построенного по схемотехнической модели рис. 2 и 3 и формулам (1)-(3), реализована в САПР AWR Microwave Office. Сначала по методике, описанной в (6), был проведен расчет идеализированной модели циркулятора без потерь на центральной частоте $f=2.4$ ГГц с параметрами $M_s=1700$ Гс, $H_0=1800$ Э, $L_0=0.48$ нГн, $C_p=3$ пФ, $L_s=3.85$ нГн, $C_s=1.1$ пФ, $L_c=1.2$ нГн, $C_c=3.6$ пФ. Далее в модель были введены: потери ($\Delta H=45^\circ$), паразитные реактивные компоненты: $R_0=0.2$ Ом, $C_{c12}=C_{c23}=C_{c31}=0.015$ пФ, $L_{iL0} = L_{iE0}=0.05L_0$, и асимметрия общего контура: $K_{as}=0.1$, $r=1$, $\varphi=0^\circ$.

На рис. 4 представлены результаты расчета частотных характеристик идеализированной модели широкополосного циркулятора и схемной модели циркулятора с паразитными элементами. Паразитные последовательные индуктивности приводят к фактическому смещению полосы рабочих частот на 100-150 МГц ниже по оси частот и к необходимости коррекции номиналов последовательной полоснорасширяющей цепи: $L_s=5.9$ нГн, $C_s=0.8$ пФ, которые подбираются по критерию максимизации полосы рабочих частот по уровню -20 дБ развязок. Потери в циркуляторе составили величину порядка 0,6 дБ в рабочей полосе частот. Можно отметить, что введение потерь приводит к различию частотных зависимостей коэффициента отражения и уровня развязок в рабочем диапазоне частот. А паразитная емкость связи приводит к вырождению «трехгорбости» частотных характеристик.

При осуществлении предельного перехода в обобщенной схеме (рис. 2, 3) с паразитными компонентами выявлено полное соответствие результатов, полученных согласно модели, асимметричного циркулятора, результатам для симметричной структуры, что подтверждает корректность предложенной в статье СВЧ теории для асимметричных циркуляторов на сосредоточенных элементах.

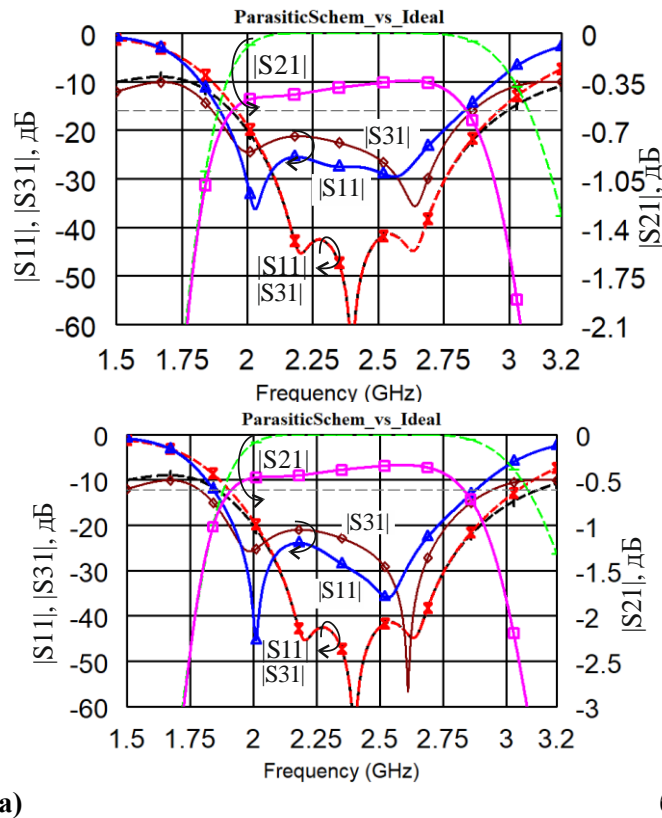


Рис. 4. Результаты расчета частотных характеристик симметричного циркулятора без паразитной емкости связи (а) и с емкостью связи (б)

На рис. 5 показаны результаты расчета полного набора элементов матрицы рассеяния циркулятора с асимметрией точки установки общего контура без паразитной емкости связи. Поскольку смещение точки установки элементов общего контура от центра симметрии ферритового элемента было сделано в направлении противоположном первому порту, то характеристики $|S22|$ и $|S33|$, $|S12|$ и $|S31|$, $|S21|$ и $|S13|$ совпадают и трансформировались одинаково, причем, на нижних частотах характеристики улучшилось, на верхних – ухудшились по сравнению с симметричным случаем. Для $|S11|$ и $|S23|$ характер изменения противоположный. Вместе с тем, существенного изменения поведения частотных характеристик не наблюдается, что было ожидаемо, исходя из физической сущности описываемых конструктивных особенностей. Симметричное добавление паразитных емкостей связи сохраняет общие тенденции в трансформации частотных характеристик, и дополнительно приводит к смещениям по частотной оси областей с пиковыми значениями характеристик согласования и развязок.

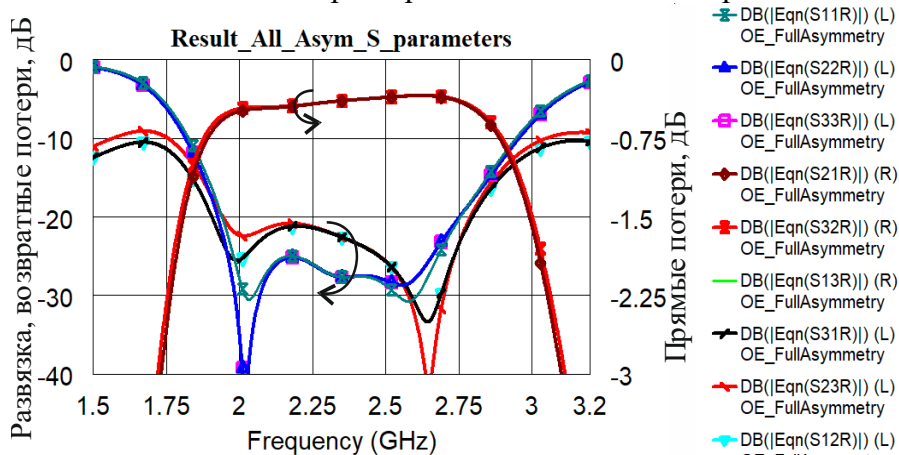


Рис. 5. Результаты расчета частотных характеристик циркулятора с асимметрией общего контура без паразитной емкости связи

Затем в схему были введены элементы асимметрии, которая ассоциируется с погрешностями изготовления топологии системы переплетения проводников. Предполагается, что при малых значениях технологических погрешностей влияния на функциональную часть циркулятора, определяемую матрицей $[S_{ij}]$, не оказывается, а изменяются номиналы паразитных реактивных элементов схемы рис.2а: L_{E1-3} , L_{L1-3} , $C_{C12/23/31}$. Можно допустить, что такие условия соблюдаются при тонкопленочной технологии изготовления топологии, когда основной вклад в асимметрию циркулятора оказывается погрешностью механического позиционирования на ферритах тонкой многослойной диэлектрической структуры с системой переплетенных проводников.

При указанных допущениях зададимся разбросом номиналов паразитных элементов схемы в виде: $L_{E/L1} = L_{E/L}(1 + \Delta_1/100)$, $L_{E/L2,3} = L_{E/L}(1 - \Delta_{2,3}/100)$, где $\Delta_{2,3} = 1 - 0.5\Delta_1 - 0.5\sqrt{4 - 3\Delta_1^2}$, а Δ_1 – отклонение номинала индуктивности в %, вызванное смещением топологии в направлении, противоположном от первого порта Y-сочленения, $L_{E/L}$ – номиналы симметричной схемы. Также положим неравномерность прижатия такой, что $C_{C12} = C_{C31}$, а $C_{C23} = C_{C12}(1 - \Delta_c/100)$. Результаты расчета частотных характеристик при $\Delta_1 = 5\%$ и $\Delta_c = 5\%$ представлены на рис.6а. На рис.6б показаны результаты расчета с компенсацией асимметрии Y-сочленения.

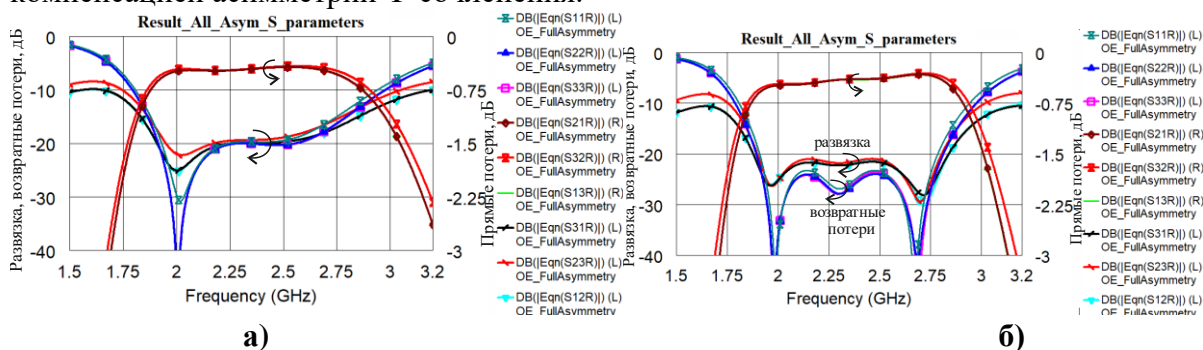


Рис. 6. Результаты расчета частотных характеристик циркулятора с асимметрией плеч (а) и с компенсацией асимметрии (б)

Компенсация асимметрии проводилась по критериям максимизации полосы рабочих частот по уровню 20 дБ развязок, минимизации сдвига центральной частоты рабочего диапазона, трехгорбого характера частотных зависимостей и близости соответствующих частотных характеристик развязок и возвратных потерь трех плеч циркулятора друг к другу. Для решения поставленной задачи, во-первых, были скорректированы номиналы элементов общего контура: $L_c = 1.7$ нГн, $C_c = 2.9$ пФ, что позволило сместить центральную частоту вверх по оси частот. Вариация номиналов согласующих цепей позволила в достаточной степени графически совместить частотные характеристики трех плеч циркулятора: $C_{p1} = 3.03$ пФ, $L_{s1} = 5.85$ нГн, $C_{s1} = 0.81$ пФ, $C_{p2} = 3.04$ пФ, $L_{s2} = 5.1$ нГн, $C_{s2} = 0.93$ пФ, $C_{p3} = 3.04$ пФ, $L_{s3} = 5.2$ нГн, $C_{s3} = 0.92$ пФ. Таким образом, запас на регулировку реактивных элементов последовательных согласующих цепей в 15 % и цепей общего контура в 30 % позволяет с достаточной для практики точностью скомпенсировать асимметрию Y-сочленения, вызванную погрешностью установки системы переплетенных проводников.

Результаты тестирования модели узкополосного вентиля с комплексной согласующей нагрузкой

Модель учета внешней асимметрии циркулятора на сосредоточенных элементах была применена при проектировании узкополосного вентиля, рассчитанного на высокий уровень мощности. Трехмерная модель устройства показана на рис.7а. Высокий уровень проходной мощности вынуждает применять большой по размерам согласующий резистор в третьем плече циркулятора. Большие габариты нагрузки R_L приводят к значительной величине номинала паразитной емкости C_L и необходимости введения компенсирующей

индуктивности L_k . Индуктивность позволяет обеспечить циркуляцию на заданной частоте, но уменьшает полосу частот развязки. Три схемных элемента R_L , C_L , L_k образуют параллельный резонансный контур в нагрузке третьего плеча циркулятора.

Результаты моделирования частотных характеристик вентиля, полученные согласно схемной модели (3) с учетом паразитных элементов конструкции, представлены на рис.7б. На рисунке также приводятся экспериментальные результаты, полученные на макетном образце устройства. Гибкость схмотехнической модели позволила получить хорошее соответствие результатов расчета и измерений. При этом прямые потери на натурном образце выше расчетных.

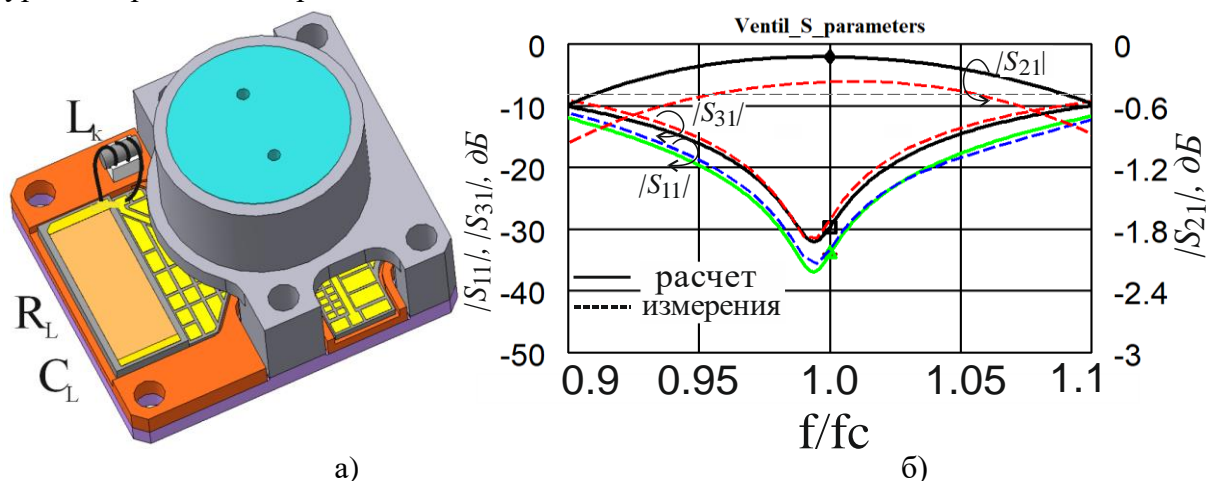


Рис. 7. S-параметры вентиля на высокий уровень мощности

Выводы

Таким образом, в ходе проведенных исследований была показана важность учета при схмотехническом проектировании широкополосных циркуляторов на сосредоточенных элементах паразитных элементов конструкции Y-соединения. Предложенная в статье новая модель асимметричного циркулятора позволила провести анализ влияния внутренней асимметрии ферритового Y-соединения на его частотные характеристики. В работе показана возможность компенсации внутренней асимметрии посредством вариации параметров элементов согласующих цепей. Моделирование показало, что для этого необходимо заложить допуск на регулировку порядка нескольких десятков процентов. Применение схемной модели узкополосного циркулятора с комплексной нагрузкой в третьем плече при проектировании вентиля, рассчитанного на высокий уровень мощности, показало высокую степень соответствия теории и практики. Предложенная модель асимметричного широкополосного циркулятора позволяет использовать её в комбинации с другими СВЧ устройствами в системе микроволнового приемопередатчика.

Библиографический список

1. Helzajn J. The Stripline Circulator Theory and Practice. *John Wiley & Sons*, 2008, 604 p.
2. Konishi Y. Lumped element Y circulator. *Transactions on Microwave Theory and Techniques*, 1965, vol. 13, no. 6, pp. 852-864.
3. Патент US 5838209 (United States). Nonreciprocal junction circuit element having different conductor intersecting angles/ Okada T., Hasegawa N., Tokudera H. // Murata Manufacturing Co., Ltd.; приоритет от 17.11.1998.
4. Патент US 7365616B2 (United States). Non-reciprocal element with three central conductors and communication apparatus using the same/ Takeda S., Mikami H., Ichikawa K.// Hitachi Metals, Ltd.; приоритет от 29.04.2008. 17 с.
5. R. Knerr. An Improved Equivalent Circuit for the Thin Film Lumped-Element Circulator. *IEEE Transactions on MTT*, Vol. MTT-20, No. 7, July 1972. pp.446-452.
6. Сорокин А.В. Проектирование малогабаритных ферритовых развязывающих устройств на сосредоточенных элементах / Диссертация на соискание ученой степени кандидата технических наук // Нижний Новгород, 2014. сс. 293.

Е.А. Mihalitsyn

CIRCUIT MODEL OF ASYMMETRIC LUMPED ELEMENT CIRCULATORS

Branch of Federal State Unitary Enterprise «Russian Federal Nuclear Center - All-Russian Scientific-Research Institute of Experimental Physics» « Federal Research and Production Center Scientific-Research Institute of Measuring Systems named after Yu.E. Sedakov»

The research objects are the ferrite lumped elements circulators having branches asymmetry. The article is concerned to schematic analysis of narrow and wide band asymmetric lumped element ferrite isolators and circulators. Two asymmetry types of circulators are considered in the paper, the first is provided with difference in circulator branches load impedances and the second is defined with technology specific of meshed conductors in ferrite Y-junction. Ferrite losses are considered in the model. Schematic design methodology is integrated to the AWR Microwave Office high frequency devices CAD system.

Keywords: asymmetric lumped element circulator, meshed conductor's system, spurious circuit elements.

МАЛОГАБАРИТНЫЕ ПОГЛОЩАЮЩИЕ ЭЛЕМЕНТЫ ДЛЯ РЕАЛИЗАЦИИ СРЕДНИХ И БОЛЬШИХ ОСЛАБЛЕНИЙ

Нижегородский государственный технический университет им. Р.Е. Алексеева

Предложены топологии поглощающих чип-элементов с однородной резистивной пленкой, обеспечивающие в малых габаритах реализацию средних и больших ослаблений при снижении в среднем в 2 – 3 раза максимальных значений градиента потенциала и плотности мощности по сравнению с типовыми топологиями путем расширения на некоторой длине заземленных контактов и оптимизации профиля входного (выходного) контакта. Полученные результаты существенны для работы поглощающих элементов в импульсном режиме.

Ключевые слова: резистивные пленочные поглощающие элементы, фиксированные аттенюаторы.

Введение

Широкополосные фиксированные аттенюаторы с пленочными поглощающими элементами (ПЭ) широко используются в телевизионных, радиоприемных и измерительных системах [1]. ПЭ таких аттенюаторов выполняются, как правило, на основе распределенных резистивных структур, отличающихся от ПЭ с дискретными пленочными резисторами (по П- или Т- схемам) повышенной широкополосностью, малой чувствительностью к точечным дефектам и неоднородностям, разбросу технологических параметров, большей надежностью и устойчивостью к температурным воздействиям [2, 3].

Для реализации требуемого диапазона ослаблений используются ПЭ прямоугольной (большие и средние ослабления) и Т-образной топологии (средние и малые ослабления) на основе однородной резистивной плёнки [4].

Уменьшение габаритов ПЭ приводит в рамках указанных топологий к затруднениям в реализации как малых, так и больших ослаблений и требует перехода к ПЭ на основе кусочно-однородных структур [5]. Такие ПЭ отличаются дополнительной резистивной плёнкой повышенного (большие ослабления) или пониженного (малые ослабления) удельного поверхностного сопротивления между входным и выходным контактами, что усложняет процесс изготовления, понижает температурную стабильность, ухудшает работу в импульсном режиме работы и, как правило, требует введения дополнительной операции подгонки.

Цель работы – предложить топологии малогабаритных ПЭ на основе однородной резистивной пленки для реализации средних и больших ослаблений с существенно меньшими значениями градиентов потенциалов и плотностей мощности в резистивной пленке (за счёт расширения на некоторой длине заземленных контактов и оптимизации профиля входного и выходного контактов) для улучшения работы ПЭ в импульсном режиме работы.

Предлагаемая топология ПЭ и методика расчета

Предлагаемая топология приведена на рис. 1 (1, 3 и 2 – входной, выходной и заземленный контакты, 4 – резистивная пленка).

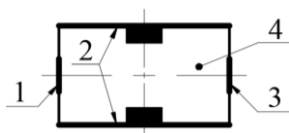


Рис. 1.

Половинная (относительно оси симметрии топологии рис. 1) структура предлагаемой топологии показана на рис. 2,а. Отметим, что ПЭ рис. 1 обладает одинаковым со структурой рис. 2,а ослаблением и при одинаковом удельном сопротивлении резистивной плёнки ρ имеет уменьшенное вдвое входное сопротивление.

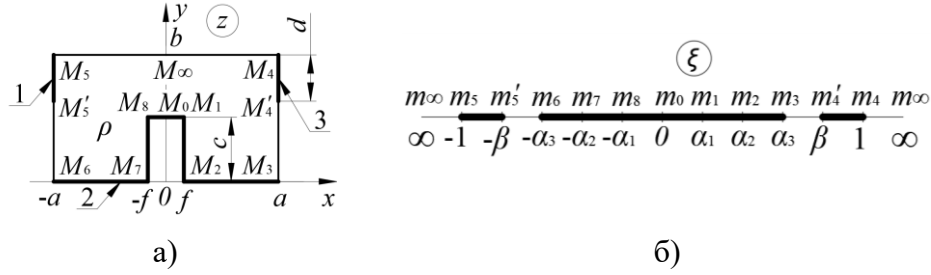


Рис. 2. Половинная структура ПЭ рис. 1 в плоскости Z (а) и её отображение в верхней полуплоскости $\xi \geq 0$ (б): 1, 3 и 2 – входной, выходной и заземленный контакты

Отображаем верхнюю полуплоскость $\xi \geq 0$ с указанным на рис. 2,б соответствием точек на внутренность многоугольника $M_0M_1M_2M_3M_4M_5M_6M_7M_8$ в плоскости Z с помощью интеграла Кристоффеля-Шварца [6]:

$$z = A \int_0^{\xi} \sqrt{(\xi^2 - a_1^2)/(\xi^2 - a_2^2)(\xi^2 - a_3^2)(\xi^2 - 1)} d\xi, \quad (1)$$

где A – масштабный коэффициент, не существенный в задаче; a_i – константы отображения.

Сопротивления П-образной схемы замещения R_{13} и $R_{12} = R_{23}$ между контактами 1 и 3, 1 и 2, а также входное сопротивление R_T и ослабление q_T (дБ) согласованного ПЭ определяются соотношениями:

$$R_{13}/\rho = 2[K'(k)/K(k) - K[(k_1)/K(k_1)], \quad R_{12}/\rho = K'[(k_1)/K(k_1)], \quad (2)$$

$$k_1^2 = (a_3/\beta)^2(1 - \beta^2)/(1 - a_3^2), \quad k^2 = (\beta - a_3)(1 - \beta + 2a_3)/(\beta + a_3)(1 - \beta), \quad (3)$$

$$R_T = R_{12}/[1 + 2/(R_{13}/R_{12})]^{1/2}, \quad q_T = 20 \lg \{ 1 + R_{13}/R_{12} + [(R_{13}/R_{12})^2 + 2(R_{13}/R_{12})]^{1/2} \}, \quad (4)$$

где $K(r)$, $K'(r)$ – полные эллиптические интегралы 1-го рода модуля r и дополнительного модуля $r' = (1 - r^2)^{1/2}$ соответственно.

По заданным величинам R_T , q_T определяем R_{13}/ρ , R_{12}/ρ , а по ним значения k_1^2 , k^2 . По известным значениям k_1^2 , k^2 находятся параметры a_3 и β .

Значения констант a_1 и a_2 находятся в соответствии с (1) из системы уравнений

$$f/c = I_1/I_2, \quad (b - d)/b = I_3/I_4, \quad (5)$$

$$I_1 = A \int_0^{a_1} \sqrt{(a_1^2 - \xi^2)/(a_2^2 - \xi^2)(a_3^2 - \xi^2)(1 - \xi^2)} d\xi, \quad (6)$$

$$I_2 = A \int_{a_1}^{a_2} \sqrt{(\xi^2 - a_1^2)/(a_2^2 - \xi^2)(a_3^2 - \xi^2)(1 - \xi^2)} d\xi, \quad (7)$$

$$I_3 = A \int_{a_3}^{\beta} \sqrt{(\xi^2 - a_1^2)/(a_2^2 - \xi^2)(\xi^2 - a_3^2)(1 - \xi^2)} d\xi, \quad (8)$$

$$I_4 = A \int_{a_3}^1 \sqrt{(\xi^2 - a_1^2)/(a_2^2 - \xi^2)(\xi^2 - a_3^2)(1 - \xi^2)} d\xi, \quad (9)$$

Гиперэллиптические интегралы $I_1 - I_4$ вычисляются с использованием квадратурных формул наивысшей алгебраической степени точности [7] после приведения их интервалов интегрирования к стандартному $(-1, 1)$ и выделения особенностей в весовую функцию с помощью подстановки $\xi = x(a_v - a_n)/2 + (a_v + a_n)/2$, где a_v , a_n –

верхний и нижний пределы интегрирования. При этом каждый интеграл приводится к виду

$$\int_{-1}^1 \frac{g(x)dx}{\sqrt{1-x^2}} = \frac{\pi}{N} \sum_{n=1}^N g\left(\cos \frac{2n-1}{2N} \pi\right) + R_N(g)$$

где остаточный член $R_N(g) = (\pi/N)(g^{2N}(\eta)/2N!)$, $-1 < \eta < 1$.

После определения констант a_1 и a_2 с помощью любого из интегралов (6) – (9) ($I_1 = f$, $I_2 = c$, $I_3 = b - d$, $I_4 = b$) находится координата у точки M_4' (M_5') выходного (входного) контактов из равенства $I_3 = b - d$.

Результаты расчетов симметричного ПЭ рис. 2,а приведены на рис. 3.

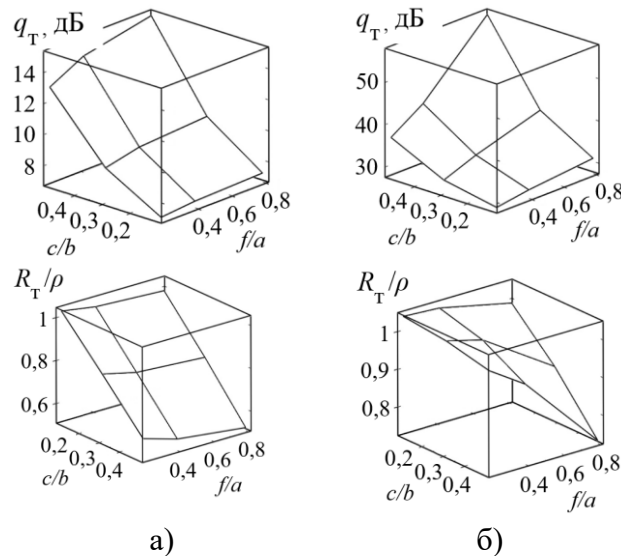


Рис. 3. Зависимость ослабления и нормированного входного сопротивления симметричного ПЭ рис. 2,а при $b/a = 5$ (а) и $b/a = 1$ (б)

Задачу можно решить и по-другому. После определения константы a_3 любые значения a_1 и a_2 из интервала $0 < a_1 < a_2 < a_3$ позволяют по соотношениям (6) – (9) получить множество ПЭ с заданными значениями R_T/ρ и q_T , отличающихся размерами c и $2f$. Этот удивительный результат невозможно получить никаким другим способом.

В частном случае $c = f = 0$ ($a_1 = a_2 = 0$) интеграл Кристоффеля-Шварца (1) вырождается в эллиптический интеграл 1-го рода в форме Лежандра, обращением которого является эллиптический синус Якоби. Последний осуществляет обратное отображение – многоугольника в плоскости Z на полуплоскость $\xi \geq 0$. При этом соотношения (2) – (4) переходят в известные для прямоугольного ПЭ:

$$R_T/\rho = \{K'(n)K(n_1)/[K(n)K'(n_1)]\}^{1/2}, \quad (10)$$

$$q_T = 20 \lg\{1+2/(A-1)\}, \quad A = [K'(n_1)K'(n)/(K(n_1)K(n))]^{1/2}, \quad (11)$$

$$n = cn(K'd/b, s'), \quad n_1 = cn[(K'(1-d/b), s'), \quad K'(s)/K(s) = b/a. \quad (12)$$

Здесь использованы обозначения, принятые в теории эллиптических интегралов и эллиптических функций Якоби [6].

Топология ПЭ с оптимизированным профилем контактов

Сильная неравномерность распределения плотности мощности и градиента потенциала по поверхности ПЭ (наибольшие значения перечисленные параметры достигают вблизи входного контакта) может привести, особенно в импульсном режиме, к выгоранию резистивной плёнки вблизи входного контакта ПЭ. Снижение градиента потенциала и плотности мощности на входе ПЭ требует оптимизации профиля

входного(выходного) контакта [8]. Несимметричный ПЭ с профилем входных контактов, близким к оптимальному, представлен на рис. 4.

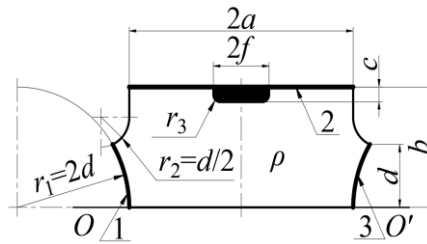


Рис. 4. ПЭ со скругленным профилем контактов

Расчеты проводились в программном комплексе *Elcut*.

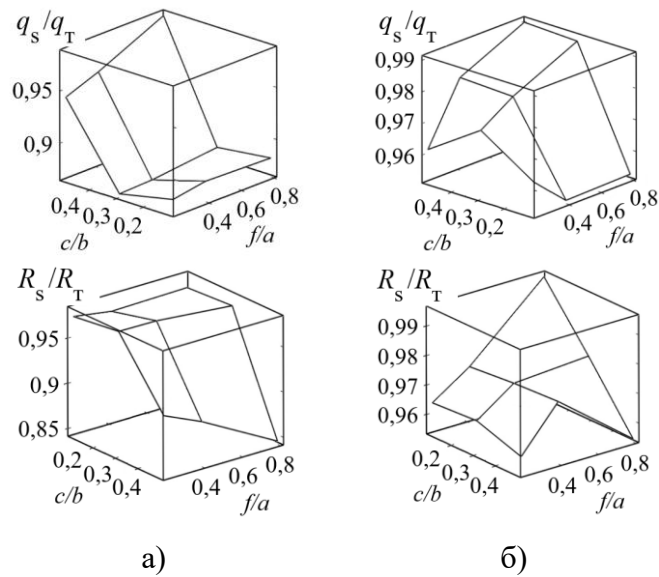


Рис. 5. Зависимость отношения ослаблений (q_s/q_T) и входных сопротивлений (R_s/R_T) ПЭ скругленной и прямоугольной конфигураций от c/b и f/a при $b/a = 5$ (а) и $b/a = 1$ (б)

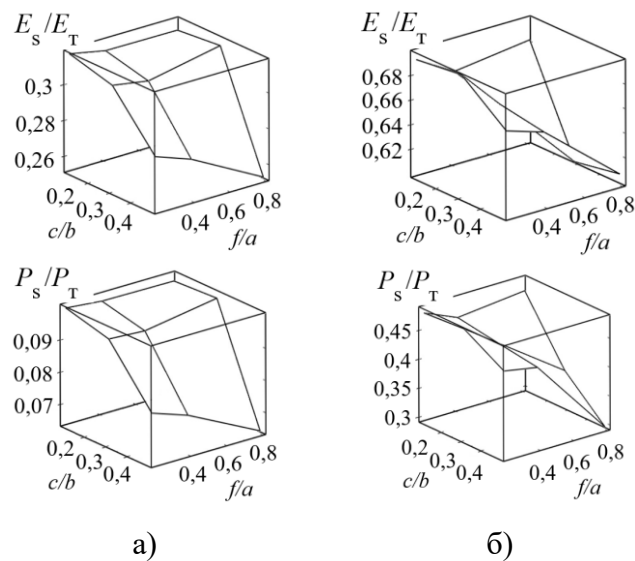


Рис. 6. Зависимость отношения градиентов потенциала (E_s/E_T) и плотностей мощности (P_s/P_T) вблизи входного контакта ПЭ скругленной и прямоугольной конфигурации от c/b и f/a при $b/a = 5$ (а) и $b/a = 1$ (б)

Результаты расчетов показывают (рис. 5), что изменение ослабления и входного сопротивления при переходе от прямоугольной к скругленной топологии ПЭ в диапазоне $5 \leq b/a \leq 1$, $0,1 \leq c/b \leq 0,5$, $0,3 \leq f/a \leq 0,9$ не превосходит 10 %.

Уменьшение градиента потенциала и плотности мощности при таком переходе (рис. 6) сильнее проявляется при больших b/a . При $b/a = 5$ градиент потенциала уменьшается в области входного контакта в 3 раза, а плотность мощности в 15 раз. При $b/a = 1$ получаем соответственно 1,7 и 3,3 раза. Отношение $R_{вх}/\rho$ мало меняется от b/a и составляет в среднем 0,4 – 0,45.

Аппроксимацией результатов моделирования получены расчетные соотношения с погрешностью менее 3 % .

Выводы

Предложены топологии малогабаритных ПЭ на основе однородной резистивной пленки, позволяющие реализовать широкий диапазон ослаблений при снижении в среднем в 2 – 3 раза градиента потенциала и плотности мощности вблизи входного контакта, что важно для работы в импульсном режиме.

Библиографический список

1. Львов А.Е., Рудоясова Л.Г., Шишков Г.И., Щербаков В.В. Тонкопленочные полосковые аттенюаторы // Труды НГТУ им. Р.Е. Алексеева, 2014, № 5, с. 40–46.
2. Чижиков С.В. Анализ тенденций развития чип-резистивных СВЧ-аттенюаторов // Успехи современной радиоэлектроники, 2017, № 8, с. 36–42.
3. Бажилов В.А., Михалицын Е.А. Проектирование малогабаритных аттенюаторов с малым уровнем ослабления // Физика волновых процессов и радиотехнические системы, 2014, т. 17, № 1, с. 54–63.
5. Уткин В.Н., Садков В.Д., Якимов Д.Ю. Расчет Т-образных поглощающих элементов чип-аттенюаторов для поверхностного монтажа // Известия вузов. Электроника, 2008, № 1, с. 86–88.
6. Sadkov, V., Fomina, K. & Pilkevich, A. Absorbing Elements Based on a Uniform Resistive Film for the Implementation of a Wide Range of Radio Signal Power Attenuations. Russian Microelectronics v. 48, № 7, 2019, pp. 516–521.
7. Лаврентьев М.А., Шабат Б.В. Методы теории функций комплексного переменного. - М.: Гл. ред. физ. - мат. лит., 1973. - 749 с.
8. Крылов В.И. Приближенное вычисление интегралов. - М.: Наука, 1967. - 500 с.
9. Sadkov, V.D., Utkin, V.N. Optimal profile of input contacts of the attenuator plate on the basis of distributed resistive structures. Radioelectronics and Communications Systems v. 51, № 12, 2008, pp. 675–677.

A.V. Pilkevich, V.D. Sadkov

**SMALL ABSORBING ELEMENTS FOR THE IMPLEMENTATION
OF MEDIUM AND LARGE IMPAIRMENTS**

Nizhny Novgorod state technical University n.a. R.E. Alekseeva

Purpose: The article proposes topologies of absorbing elements with uniform resistive film for waveguide, coaxial and strip attenuators.

Design/methodology/approach: Optimization of absorbing elements is performed by extending the grounded contacts at a certain length and optimizing the profile of the input (output) contact.

Findings: Graphs are provided showing the range of attenuations implemented, the gain in value of potential gradient and power density depending on the ratio of dimensions in the absorbing element.

Research limitations/implications: Size restrictions dictate the accuracy of the equipment.

Originality/value: The proposed topology of the absorbing element shows a decrease in the maximum values of the potential gradient and power density in comparison with the typical topology by an average of 2 – 3 times.

Keywords: resistive film absorbing elements, fixed attenuators, potential gradient, power density, optimal profile.

МАТЕМАТИЧЕСКИЙ АППАРАТ ДЛЯ МОДЕЛИРОВАНИЯ ПАССИВНЫХ ЭЛЕМЕНТОВ ИНТЕГРАЛЬНЫХ СХЕМ

Нижегородский государственный технический университет им. Р.Е. Алексеева

Обобщена методика решения электрических, электростатических, магнитостатических, тепловых и электродинамических задач, возникающих при моделировании пленочных и полупроводниковых пассивных элементов интегральных схем с использованием формулы Келдыша – Седова для поля потенциалов.

Ключевые слова: стационарные и плоские электродинамические задачи, формула Келдыша – Седова.

Обобщена методика решения электрических, электростатических, магнитостатических, тепловых и электродинамических задач, возникающих при проектировании пленочных и полупроводниковых пассивных элементов интегральных схем. Расчет таких элементов сводится, как правило, к решению двумерных смешанных краевых задач для уравнения Лапласа и волнового уравнения, которые методами конформных отображений приводятся к соответствующим задачам для верхней полуплоскости.

Статическая задача формулируется следующим образом: на действительной оси даны раздельно лежащие отрезки (a_k, b_k) , $k = 1, 2, \dots, n$ ($a_1 < b_1 < \dots < b_n$). Через D' обозначена совокупность этих отрезков, через D'' – остальная часть действительной оси ($D = D' + D''$ – вся действительная ось), так что D'' состоит из конечного числа отрезков (b_k, a_{k+1}) ($k = 1, 2, \dots, n-1$) и из бесконечного отрезка (b_n, a_1) , состоящего из полупрямых $b_n < x < \infty$ и $-\infty < x < a_1$. Требуется найти функцию $\Phi(z) = U+iV$, гомоморфную в верхней полуплоскости и ограниченную на бесконечности, по граничному условию $U^+ = f(t)$ на D' и $V^+ = g(t)$ на D'' .

Известно, что задача не имеет решений, ограниченных вблизи всех концов отрезков. Если отказаться от условия ограниченности $\Phi(z)$ и требовать лишь ограничения интеграла от нее (функции $f(t)$ и $g(t)$ удовлетворяют условию Гельдера), то задача решается с точностью до $n+1$ произвольного постоянного. Наконец, задача имеет единственное решение, если, кроме того, потребовать, чтобы $\Phi(z)$ была ограниченной вблизи каких-либо отрезков и задать ее значение в некоторой точке. Тогда общее решение задачи класса h_0 имеет вид [1]:

$$\Phi(z) = \Psi(z) + \varphi_0(z), \quad (1)$$

где $\varphi_0(z)$ – общее решение класса h_0 однородной задачи (при $f(t) = g(t) = 0$),

$$\varphi_0(z) = \frac{C_0 z^n + C_1 z^{n-1} + \dots + C_n}{\sqrt{\prod_{k=1}^n (z - a_k)(z - b_k)}}$$

C_0, \dots, C_n – произвольные действительные числа;

$\Psi(z)$ – одно из частных решений исходной задачи,

$$\Psi(z) = \frac{1}{\pi i} \sqrt{\frac{\prod_{k=1}^n (z - a_k)}{\prod_{k=1}^n (z - b_k)}} \int_D \frac{\prod_{k=1}^n (z - b_k)}{\prod_{k=1}^n (z - a_k)} \frac{h(t) dt}{t - z}, \quad h(t) = \begin{cases} f(t) \text{ на } D^I \\ ig(t) \text{ на } D^{II} \end{cases}$$

Общий вид решения исходной задачи класса $h(a_1, \dots, a_n)$ имеет вид:

$$\Phi(z) = \psi(z) + C \sqrt{\prod_{k=1}^n (z - a_k) / \prod_{k=1}^n (z - b_k)},$$

где C – произвольная постоянная.

Решение $\varphi_0(z)$ описывает структуры, в которых участки (a_k, b_k) эквипотенциальны (на них $f(t) = \text{const}$), а на (b_k, a_{k+1}) задано условие $q(t) = 0$ (граница структуры – линия тока).

Искомая потенциальная функция определяется как $U(x,y) = \text{Re}\Phi(z)$. По известной потенциальной функции матрица, например, погонных емкостей $[C]$ задачи (с учетом электрогидродинамических аналогий) вычисляется следующим образом [2]:

$$C_{k,1} = 2\varepsilon X_{k,1}, C_{k,n} = 2\varepsilon X_{k,n-1}, C_{k,p} = 2\varepsilon(X_{k,p} - X_{k,p-1}), X_{k,j} = \det[E_{k,j}] / \det[F], \quad (2)$$

где ε – диэлектрическая проницаемость среды; $k = 1, 2, 3, \dots, n$; $p = 2, 3, \dots, n-1$; $j = 1, 2, 3, \dots, n$.

Элементы матриц $[F]$ и $[E_{k,j}]$:

$$[F] = \begin{bmatrix} F_{1,n-2} & F_{1,n-3} & \cdot & \cdot & F_{1,0} \\ F_{2,n-2} & F_{2,n-3} & \cdot & \cdot & F_{2,0} \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ F_{n-1,n-2} & F_{n-1,n-3} & \cdot & \cdot & F_{n-1,0} \end{bmatrix} \quad [E_{k,j}] = \begin{bmatrix} F_{1,n-2} & F_{1,n-3} & \cdot & \cdot & F_{1,0} \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ F_{j-1,n-2} & F_{j-1,n-3} & \cdot & \cdot & F_{j-1,0} \\ E_{k,n-2} & E_{k,n-2} & \cdot & \cdot & E_{k,0} \\ F_{j-1,n-3} & F_{j-1,n-3} & \cdot & \cdot & F_{j-1,0} \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ F_{n-1,n-2} & F_{n-1,n-3} & \cdot & \cdot & F_{n-1,0} \end{bmatrix}$$

выражаются следующими гиперэллиптическими интегралами

$$F_{k,m} = (-1)^k \int_{b_k}^{a_{k+1}} \frac{t^m dt}{\sqrt{R(t)}}, \quad E_{k,m} = (-1)^{k-1} \int_{a_k}^{b_k} \frac{t^m dt}{\sqrt{-R(t)}},$$

$$R(t) = (t - a_1)(t - b_1)(t - a_2)(t - b_2) \dots (t - a_n)(t - b_n).$$

Гиперэллиптические интегралы $F_{k,m}$, $E_{k,m}$ вычисляются с использованием квадратурных формул наивысшей алгебраической степени точности после приведения интервалов интегрирования к стандартному $(-1, 1)$ и выделению особенностей в весовую функцию $1/(1-w^2)^{1/2}$ [3]:

$$\int_{-1}^1 \frac{g(w)dw}{\sqrt{1-w^2}} \cong \frac{\pi}{N} \sum_{n=1}^N g\left(\cos \frac{2n-1}{2N} \pi\right).$$

При неэквипотенциальности контактов (когда удельное сопротивление материала контакта по порядку величины сопоставимо с удельным сопротивлением материала подложки, в тепловых задачах) распределение потенциала на них задается функцией $f(t)$, результирующее решение дается функцией $F(z)$, а потенциальная функция определяется как $U(x,y) = \text{Re} F(z)$.

В этом случае краевые условия имеют следующий вид:

$$\frac{du}{dn} = f(\ell) \text{ на } (a_k, b_k) \text{ и } \frac{du}{dn} = 0 \text{ на } (b_k, a_{k+1}), \quad (3)$$

где ℓ – линия границы.

Очевидное условие $\sum_{k=1}^n \int_{a_k}^{b_k} f(\ell) d\ell = 0$ выражает свойство интеграла от нормальной производной гармонической функции [4].

Поставленную задачу сводим к задаче Неймана для круга. Отобразим верхнюю полуплоскость Z на круг $\zeta = \rho e^{i\varphi}$ с помощью функции $\xi = (z - i\beta)/(z + i\beta)$, $\beta \geq 0$. При этом отрезки (a_k, b_k) переходят в дуги окружности $\rho = 1$, $\varphi_{ki} \leq \varphi \leq \varphi_{ki+1}$, а для угла φ получаем

$$\begin{cases} \sin \varphi = -2\beta x / (x^2 + \beta^2) \\ \cos \varphi = (x^2 - \beta^2) / (x^2 + \beta^2) \end{cases}, a_k \leq x \leq b_k, k = 1, 2, \dots, n. \quad (4)$$

Считая, что все точки a_k, b_k удовлетворяют неравенству

$$|x| > 1/\mu, \quad (5)$$

выбираем число β так, чтобы углы $\varphi_{ki} \leq \varphi \leq \varphi_{ki+1}$ целиком попали в два интервала

$$0 \leq \varphi \leq \pi/4 \text{ и } 7\pi/4 \leq \varphi \leq 2\pi. \quad (6)$$

Для этого, как следует из соотношений (4) - (6) нужно, чтобы $2\beta\mu/[1+(\beta\mu)^2] \leq \sqrt{2}/2$, то есть $\beta\mu < 1$. Тогда решение задачи Неймана для функции $U(\zeta)$ дается формулой Дини [4]:

$$U(\xi) = -\frac{1}{\pi} \int_0^{2\pi} f(t) \ln(e^{it} - \xi) dt + C, \quad (7)$$

где C - произвольная постоянная, которую определяем при дополнительном условии $U(\xi_0) = U(e^{i\varphi_0}) = U_0$ - известное значение функции $U(\zeta)$ в некоторой точке ζ_0 границы.

Равенство (7) запишем в виде

$$C = U_0 + \frac{1}{\pi} \int_0^{2\pi} f(t) \ln(e^{it} - e^{i\varphi_0}) dt,$$

а, вводя новую функцию $\bar{U}(\xi) = U(\xi) - U_0$, получим

$$\bar{U}(\xi) = -\frac{1}{\pi} \int_0^{2\pi} f(t) \ln \frac{|e^{it} - \xi|}{|e^{it} - e^{i\varphi_0}|} dt.$$

Для значений этой функции на границе получаем

$$\bar{U}(e^{i\varphi}) = -\frac{1}{\pi} \int_0^{2\pi} f(t) \ln \frac{|e^{it} - \xi|}{|e^{it} - e^{i\varphi_0}|} dt = -\frac{1}{\pi} \sum_{k=1}^n \int_{\varphi_{ki}}^{\varphi_{ki+1}} f(t) \ln \frac{|\sin[(t-\varphi)/2]|}{|\sin[(t-\varphi_0)/2]|} dt.$$

Введем проводимости по усредненным по проводникам потенциалам и токам:

$$(\varphi_{mi+1} - \varphi_{mi}) U_m^{cp} = \int_{\varphi_{mi}}^{\varphi_{mi+1}} \bar{U}(\varphi) d\varphi = -\frac{1}{\pi} \sum_{k=1}^n \int_{\varphi_{ki}}^{\varphi_{ki+1}} f(t) \left[\int_{\varphi_{mi}}^{\varphi_{mi+1}} \ln \frac{|\sin[(t-\varphi)/2]|}{|\sin[(t-\varphi_0)/2]|} d\varphi \right] dt.$$

Обозначив

$$S_m(t) = \int_{\varphi_{mi}}^{\varphi_{mi+1}} \ln \frac{|\sin[(t-\varphi)/2]|}{|\sin[(t-\varphi_0)/2]|} d\varphi,$$

получим

$$(\varphi_{mi+1} - \varphi_{mi})U_m^{\text{cp}} = -\frac{1}{\pi} \sum_{k=1}^n \int_{\varphi_{ki}}^{\varphi_{ki+1}} f(t) S_m(t) dt. \quad (8)$$

Используя обобщенную теорему о среднем для (8) с учетом выполнения условия знакопостоянства функции $S_m(t)$ по (6) запишем

$$(\varphi_{mi+1} - \varphi_{mi})U_m^{\text{cp}} = -\frac{1}{\pi} \sum_{k=1}^n \sigma f(\varphi_k) (\varphi_{k+1} - \varphi_k) \frac{\int_{\varphi_{ki}}^{\varphi_{ki+1}} S_m(t) dt}{\sigma(\varphi_{k+1} - \varphi_k)},$$

где $\sigma = 1/\rho$ – удельная поверхностная проводимость среды в области S .

Величина $\sigma(\varphi_{k+1} - \varphi_{ki})f(\varphi_k) = I_k^{\text{cp}}$ – средний ток через дугу на границе $\rho = 1$, $\varphi_{ki} \leq \varphi \leq \varphi_{ki+1}$, среднее напряжение на m -ной дуге дается формулой

$$U_m^{\text{cp}} = \sum_{k=1}^n I_k^{\text{cp}} R_{mk},$$

где выражения для взаимных сопротивлений в системе проводников имеет вид:

$$R_{km} = \frac{1}{\pi} \frac{\int_{\varphi_{ki}}^{\varphi_{ki+1}} S_m(t) dt}{\sigma(\varphi_{k+1} - \varphi_k)}.$$

Таким образом, находим параметры многополюсника (n -полюсника) или, что тоже самое, матрицу сопротивлений исходной задачи.

Аналогичным образом решается краевая задача

$$\frac{du}{dn} = f(\ell) \text{ на } (a_k, b_k) \text{ и } \frac{du}{dn} = g(\ell) \text{ на } (b_k, a_{k+1}).$$

Методы теории функций комплексного переменного эффективны и при решении плоских задач электродинамики. Они позволяют свести задачу к решению интегральных уравнений для тангенциальных составляющих токов вдоль границы. Это могут быть задачи, связанные с исследованиями микроэлектронных структур СВЧ, содержащих плоские проводники (линии передачи, фильтры, многоэлементные решетки излучателей, меандровые замедляющие системы).

Известно, что в двумерном случае электромагнитные волны могут быть электрические ($H_z = 0$) и магнитные ($E_z = 0$), а векторный потенциал заменяется скалярной функцией u , убывающей на бесконечности как $O(x^2 + y^2)^{-1/2}$ и удовлетворяющей двумерному волновому уравнению (уравнению Гельмгольца):

$$\frac{d^2 u}{dx^2} + \frac{d^2 u}{dy^2} + \gamma^2 u = 0, \quad \gamma^2 = k^2 - h^2, \quad (9)$$

где $k = \omega/c$, c – скорость света; h – постоянная распространения вдоль оси z .

В частности, для задачи о распространении волны электрического типа в системе n бесконечно протяженных проводников на вещественной оси с координатами начал и концов $-\infty < a_1 < b_1 < a_2 < b_2 < \dots < a_n < b_n < +\infty$. Это краевая задача в канонической n -связной области S с граничным условием $u = 0$ к плоскости перпендикулярного сечения системы.

После замены переменных

$$\xi = x + iy, \quad \bar{\xi} = x - iy, \quad x = (\xi + \bar{\xi})/2, \quad y = (\xi - \bar{\xi})/2, \quad v(\xi, \bar{\xi}) = d^2 u(\xi, \bar{\xi})/d\xi d\bar{\xi}$$

уравнение (9) приводится к уравнению Вольтера второго рода в комплексной области $S_1 = \{\xi, \bar{\xi}; (x, y) \in S\}$:

$$v(\xi, \bar{\xi}) = v_0(\xi) - \frac{\gamma^2}{4} \int_0^{\bar{\xi}} \int_0^{\xi} v(t, \tau) dt d\tau.$$

Его решение представимо в следующем виде [5]:

$$u(x, y) = \operatorname{Re} \int_0^{\xi} J_0(\gamma \sqrt{\xi(\xi-t)}) \frac{P_{n-2}(t)}{\sqrt{R(t)}} dt + A_n, \quad R(t) = (t-a_1)(t-b_1)(t-a_2)\dots(t-b_n), \quad (9)$$

где A_n – вещественная константа; P_{n-2} – полином степени $n-2$ с вещественными константами A_1, A_2, \dots, A_{n-1} вида $\varphi_0(z)$.

При любых вещественных значениях коэффициентов A_1, A_2, \dots, A_{n-1} функция (9) на границе области принимает значения

$$u(x, y) = C_j = \operatorname{const}, \quad j = 1, 2, \dots, n-2; \quad C_j = \operatorname{Re} \int_0^{a_j} J_0(\gamma \sqrt{\xi(\xi-t)}) \frac{P_{n-2}(t)}{\sqrt{R(t)}} dt + A_n. \quad (10)$$

При $\gamma = 0$ однородной задаче и граничному условию удовлетворяет функция

$$u(x, y) = \operatorname{Re} \left[P_{n-1}(\xi) / \sqrt{R(\xi)} \right],$$

где P_{n-1} – полином степени $n-1$ с вещественными константами.

При $\gamma \neq 0$ решение дается формулой (9) для значений γ , обращающих в 0 определитель системы (10).

Результаты решения краевой задачи в канонической области – верхней полуплоскости обобщаются на случай произвольной двумерной области путем конформного отображения её на каноническую – верхнюю полуплоскость.

Библиографический список

1. Мусхелишвили Н.И. Сингулярные интегральные уравнения. Граничные задачи теории функций и некоторые их приложения к математической физике. – М.: Наука, 1968. – 513 с.
2. Ямпурин Н.П., Широков Л.В., Садков В.Д., Потехин В.А. Современные проблемы радиоэлектроники с позиций теории конформных отображений. – Арзамас: АГПИ, 2014. – 209 с.
3. Крылов В.И. Приближенное вычисление интегралов. – М.: Наука, 1967. – 500 с.
4. Лаврентьев М.А., Шабат В.В. Методы теории функций комплексного переменного. – М.: Наука, 1973. – 736 с.
5. Фадеева Л.Б. Электромагнитные волны в СВЧ микросхемах // Электронная техника. Сер. Микроэлектроника, 1969, вып. 5, с. 19 - 26.

A.V. Pilkevich, V.D. Sadkov

MATHEMATICAL APPARATUS FOR MODELING PASSIVE ELEMENTS OF INTEGRATED CIRCUITS

Nizhny Novgorod state technical University n.a. R.E. Alekseeva

Summarized the method of solving electric, electrostatic, magnetostatic, thermal and electrodynamic problems encountered in the design of the passive film and the semiconductor elements of integrated circuits using the formulas of Keldysh – Sedov for the field potentials.

Keywords: stationary physical fields, the formula of the Keldysh – Sedov.

ТОНКОПЛЕНОЧНЫЕ ПОГЛОЩАЮЩИЕ ЭЛЕМЕНТЫ АДАПТОАТТЕНУАТОРОВ ДЛЯ ГИС

Нижегородского государственного технического университета им. Р.Е. Алексеева

Предложены перспективные топологии и методики расчета поглощающих элементов адаптоаттенуаторов – пленочных поглощающих элементов с разными значениями входного и выходного сопротивлений на основе однородных и кусочно-однородных резистивных структур для реализации широкого диапазона ослаблений с заданными значениями входного и выходного сопротивлений для гибридных интегральных схем.

Рассмотрены топологии поглощающих элементов на основе однородной резистивной пленки с проводящей или непроводящей областью, позволяющие в малых габаритах обеспечить требуемый диапазон ослаблений и коэффициентов трансформации.

Исследованы топологии на основе кусочно-однородной резистивной пленки, в том числе с оптимальным профилем входного и выходного контактов, позволяющие не только сократить размеры поглощающих элементов, но и в 2-3 раза уменьшить максимальные значения градиента потенциала и плотности мощности, что особенно важно для импульсного режима работы адаптоаттенуатора.

Ключевые слова: Пленочный поглощающий элемент, адаптоаттенуатор, градиент потенциала, плотность мощности, оптимальный профиль.

Введение

Широкополосные фиксированные аттенуаторы с пленочными поглощающими элементами (ПЭ) широко используются в телевизионных, радиоприемных и измерительных системах [1]. Разновидностью аттенуаторов являются адаптоаттенуаторы, отличающиеся разными значениями входного и выходного сопротивлений (обеспечивающие трансформацию сопротивлений), что особенно интересно для устройств на базе гибридных и полупроводниковых интегральных схем [2]. Анализ перспективных топологий ПЭ адаптоаттенуаторов, позволяющих в малых габаритах обеспечить необходимый диапазон ослаблений и коэффициентов трансформации и их моделирование с использованием аналитических и численных методов является целью настоящей работы.

Разработанная ранее для коаксиальных трактов [3], и предлагаемые топологии ПЭ адаптоаттенуаторов (обычно симметричных относительно оси OO') представлены на рис. 1а и рис. 1,б – рис. 1,з. ПЭ рис. 1,б и рис. 1,г получены обращением краевых условий ПЭ рис. 1,а и рис. 1,в соответственно [4]. Такие ПЭ обладают тем же ослаблением, а коэффициент трансформации $n^2 = R_{\text{вх}} / R_{\text{вых}}$ и удельное сопротивление пленки ρ_0 обращенного (n_o^2, ρ_o) и исходного ПЭ (n_u^2, ρ_u) связаны соотношениями: $n_o^2 = 1/n_u^2$ и $\rho_o = 1/\rho_u$.

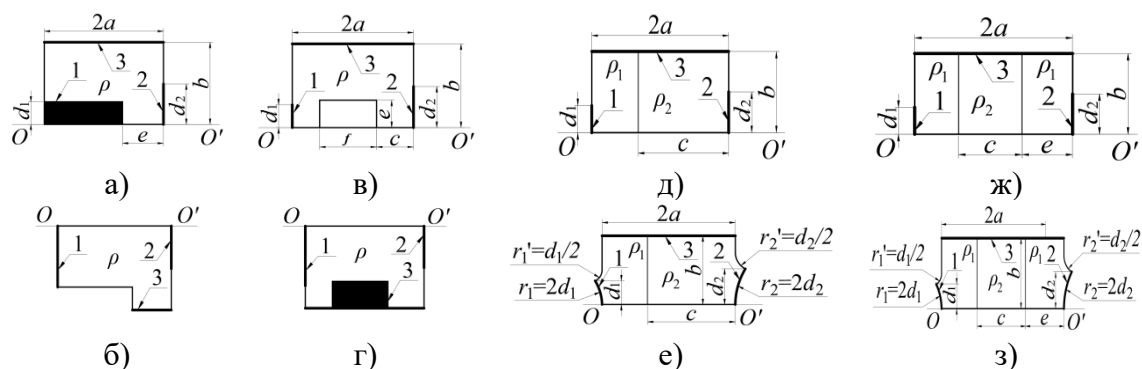


Рис. 1

Топологии ПЭ рис. 1,д – рис. 1,з позволяют за счет использования кусочно-однородной пленки обеспечить в малых габаритах реализацию всего диапазона требуемых параметров ПЭ адаптоаттенюаторов при существенном уменьшении плотности мощности и градиентов потенциала вблизи входного и выходного контактов, что особенно важно при импульсном режиме работы. Отметим, что ПЭ, симметричные относительно оси OO' обладают, по сравнению с несимметричными, теми же параметрами при увеличенном вдвое значении удельных сопротивлений резистивных пленок.

Методика расчета

Рассмотрим ПЭ рис. 1,б, представленный со всеми размерами на рис. 2,а.

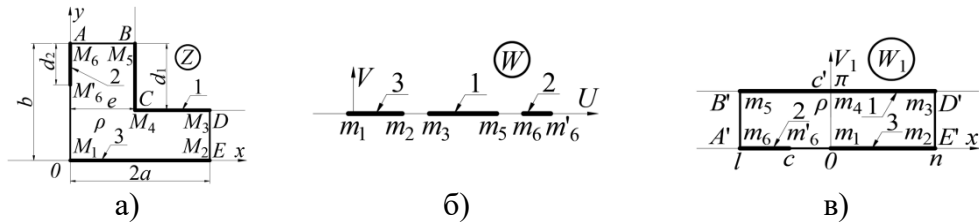


Рис. 2.

Для строгого решения задачи используем метод конформных отображений [5]. Отображаем с помощью интеграла Кристоффеля-Шварца верхнюю полуплоскость W (рис. 2,б) на внутренность шестиугольника $M_1M_2M_3M_4M_5M_6$ в плоскости Z (рис. 2,а) с указанным соответствием точек:

$$Z = A \int_0^w \frac{\sqrt{w - m_4} dw}{\sqrt{(w - 1)w(w + 1)(w - m_5)(w - m_6)}}, \quad (1)$$

где m_i – константы интеграла Кристоффеля-Шварца и выбраны $m_1 = -1, m_2 = 0, m_3 = 1$; A – масштабный коэффициент, не существенный в задаче.

Константы m_4, m_5, m_6 получаем из уравнений

$$h/b = I_1/I_2, \quad b/a = I_2/I_3, \quad 1 - h/a = I_4/I_1, \quad (2)$$

$$I_1 = A \int_{-1}^0 \frac{\sqrt{w - m_4} dw}{\sqrt{(w - 1)w(w + 1)(w - m_5)(w - m_6)}}, \quad I_2 = A \int_0^1 \frac{\sqrt{w - m_4} dw}{\sqrt{(w - 1)w(w + 1)(w - m_5)(w - m_6)}},$$

$$I_3 = A \int_{m_5}^{m_6} \frac{\sqrt{w - m_4} dw}{\sqrt{(w - 1)w(w + 1)(w - m_5)(w - m_6)}}, \quad I_4 = A \int_1^{m_4} \frac{\sqrt{w - m_4} dw}{\sqrt{(w - 1)w(w + 1)(w - m_5)(w - m_6)}}.$$

После определения констант (уравнения решаются в Mathcadе, Matlabe) вычисляется по (1) положение точки m_b при $z = m - d$ и интегрировании в пределах от m_6 до m'_6 . Дальнейший анализ структуры рис. 2,б, проводимый по [5] позволяет определить параметры схемы замещения в виде П-образной схемы с сопротивлениями R_1 и R_3 в вертикальных и R_2 – в горизонтальной ветвях и, следовательно, ослабление, входное и выходное сопротивления.

Предпочтительным является упрощенное решение задачи с контролируемой погрешностью [6]. При этом работаем не с шестиугольником $M_1M_2M_3M_4M_5M_6$, а с бесконечной полосой AM_1EDM_4Bc прямоугольным изгибом. Отображаем полосу рис. 2,в в плоскости W_1 на полосу с изгибом в плоскости Z с помощью функции

$$Z = \frac{2b}{\pi} \left[\frac{1}{\sqrt{a_1}} \arctg \beta + \frac{1}{2} \ln \frac{1 + \sqrt{a_1} \beta}{1 - \sqrt{a_1} \beta} \right], \quad \sqrt{a_1} = \frac{b}{a}, \quad \beta = \sqrt{\frac{e^w - 1}{a_1 e^w + 1}} \quad (3)$$

Для линий границы OA, OE, CD, CB в плоскости Z получаем следующие соотношения:

$$y_{OA} = \frac{a}{\pi} \left[2\sqrt{a_1} \operatorname{arctg}(\sqrt{a_1} f) - \ln \frac{1+f}{1-f} \right], \quad f = \sqrt{(1-e^{-u})/(1+a_1 e^{-u})}, \quad (4)$$

$$x_{OE} = \frac{a}{\pi} \left[2\operatorname{arctg}(g) + \sqrt{a_1} \ln \frac{1+g\sqrt{a_1}}{1-g\sqrt{a_1}} \right], \quad g = \sqrt{(e^u - 1)/(a_1 e^u + 1)}, \quad (5)$$

$$y_{CB} = \frac{a}{\pi} \left[2\sqrt{a_1} \operatorname{arctg}(\sqrt{a_1} \xi) - \ln \frac{\xi+1}{\xi-1} \right], \quad \xi = \sqrt{(1+e^{-u})/(1-a_1 e^{-u})}, \quad (6)$$

$$x_{CD} = \frac{a}{\pi} \left[2\operatorname{arctg}(\alpha) + \sqrt{a_1} \ln \frac{\alpha\sqrt{a_1} + 1}{\alpha\sqrt{a_1} - 1} \right], \quad \alpha = \sqrt{(e^u + 1)/(a_1 e^u - 1)}. \quad (7)$$

Отображение прямоугольника $t_2 m_3 m_5 m_6$ в плоскости W_1 на шестиугольник $M_2 M_3 M_4 M_5 M_6 M_1$ в плоскости Z будет строгим только при достаточном удалении линий $M_6 M_5$ и $M_2 M_3$ от начала координат. В противном случае линии $t_6 m_5$ и $t_2 m_3$ отобразятся в выпуклую линию $M_6 M_5$ и вогнутую $M_3 M_2$. Но уже при x/b и y/b больше 1,5 отношение x_{OE}/x_{CD} и y_{OA}/y_{CB} не превосходит 2 %, а погрешность определения интегральных характеристик поля потенциалов (параметров схемы замещения рис. 2, в) будет ещё меньше [6].

Для размеров n, ℓ, c и результирующего размера $n + \ell$ получаем из (4) – (7):

$$n = \left(\frac{h\pi}{a\sqrt{a_1}} \right) - \frac{2}{\sqrt{a_1}} \operatorname{arctg} \frac{1}{\sqrt{a_1}} - \ln \frac{4a_1}{1+a_1}, \quad c = \frac{\pi d}{a} + \sqrt{a_1} \operatorname{arctg} \sqrt{a_1} - \ln \frac{4a_1}{1+a_1}, \quad (8)$$

$$\ell = \frac{\pi m}{a} + \sqrt{a_1} \operatorname{arctg} \sqrt{a_1} - \ln \frac{4a_1}{1+a_1} \quad (9)$$

$$n + \ell = \frac{\pi}{a} \left(\frac{h}{\sqrt{a_1}} + \ell \right) + \left(\frac{2}{\sqrt{a_1}} + \sqrt{a_1} \right) \operatorname{arctg} \sqrt{a_1} - 2 \ln \frac{4a_1}{(1+a_1)^2}. \quad (10)$$

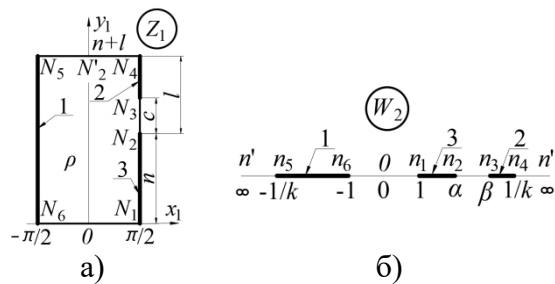


Рис. 3.

Получившийся в плоскости W_1 прямоугольник устанавливаем в плоскости Z_1 как показано на рис.3,а и отображаем на верхнюю полуплоскость W_2 (рис. 3,б) [5]:

$$w_2 = \operatorname{sn} \left(\frac{z_1}{\pi/2} K, k \right), \quad \frac{K'}{K} = \frac{n + \ell}{\pi/2},$$

$$\alpha = \operatorname{sn} \left(\frac{\pi/2 + in}{\pi/2} K, k \right) = 1 / \operatorname{dn} \left[\left(1 - \frac{\ell}{n + \ell} \right) K', k' \right]. \quad (11)$$

$$\beta = \operatorname{sn}\left(\frac{\pi/2 + i(n+c)}{\pi/2} K, k\right) = 1/\operatorname{dn}\left[\left(1 - \frac{\ell-c}{n+\ell}\right) K', k'\right]. \quad (12)$$

Здесь использованы обозначения, принятые в теории полных эллиптических интегралов 1-го рода и эллиптических функций Якоби [5].

Величины сопротивлений П-образной схемы замещения определяются с учетом (11) -(12) по соотношениям [5]:

$$R_1 = \frac{\rho}{2} \frac{\det[F]}{\det[E_{21}]}, \quad R_2 = \frac{\rho}{2} \frac{\det[F]}{\det[E_{22}]}, \quad R_3 = \frac{\rho}{2} \frac{\det[F]}{\det[E_{31}]},$$

где $\det[F]$, $\det[E_{km}]$ – определители матриц $[F]$ и $[E_{km}]$ вида

$$[F] = \begin{bmatrix} F_{11} & F_{10} \\ F_{21} & F_{20} \end{bmatrix}, \quad [E_{21}] = \begin{bmatrix} E_{21} & E_{20} \\ F_{21} & F_{20} \end{bmatrix}, \quad [E_{22}] = \begin{bmatrix} F_{11} & F_{10} \\ E_{21} & E_{20} \end{bmatrix}, \quad [E_{31}] = \begin{bmatrix} E_{31} & E_{30} \\ F_{21} & F_{20} \end{bmatrix},$$

$$F_{km} = (-1)^k \int_{b_k}^{a_{k+1}} \frac{t^m dt}{\sqrt{|R(t)|}}, \quad E_{km} = (-1)^{k+1} \int_{a_k}^{b_k} \frac{t^m dt}{\sqrt{|R(t)|}}, \quad (13)$$

$$R(t) = (t-a_1)(t-b_1)(t-a_2)(t-b_2)(t-a_3)(t-b_3), \quad a_1 = -1/k, \quad b_1 = -1, \quad a_2 = 1, \quad b_2 = \alpha, \quad a_3 = \beta, \quad b_3 = 1/k.$$

С помощью подстановок $t = x(b_k - a_k)/2 + (b_k + a_k)/2$ для E_{km} и $t = x(a_{k+1} - b_k)/2 + (a_{k+1} + b_k)/2$ для F_{km} гиперэллиптические интегралы (13) приводятся к стандартному интервалу интегрирования и вычисляются с использованием квадратурной формулы наивысшей степени точности [7]:

$$\int_{-1}^1 \frac{g(x) dx}{\sqrt{1-x^2}} = \frac{\pi}{N} \sum_{n=1}^N g\left(\cos \frac{2n-1}{2N} \pi\right) + R_N(g), \quad R_N(g) = \frac{\pi}{N} \frac{g^{2N}(\eta)}{2N!}, \quad -1 < \eta < 1.$$

После вычисления параметров $R_1 - R_3$ схемы замещения рис. 2, в коэффициент трансформации n^2 , входное, выходное сопротивления и ослабление ПЭ рис. 1,б вычисляется по известным соотношениям [8]:

$$n^2 = (R_1/R_3) (1+R_3/R_2) / (1+R_1/R_2),$$

$$R_{\text{вх}} = R_1 \sqrt{\frac{1+R_3/R_2}{1+R_1/R_2}} \sqrt{\frac{1}{1+R_1/R_2+R_3/R_2}}, \quad R_{\text{вых}} = R_3 \sqrt{\frac{1+R_1/R_2}{1+R_3/R_2}} \sqrt{\frac{1}{1+R_1/R_2+R_3/R_2}},$$

$$q = 20 \lg \left[\sqrt{(1+R_2/R_3)(1+R_2/R_1)} + \sqrt{(1+R_2/R_3)(1+R_2/R_1)-1} \right], \text{ дБ},$$

или

$$q = 20 \lg \left[(1+R_2/R_1) \sqrt{n^2} + \sqrt{n^2(1+R_2/R_1)^2 - 1} \right], \text{ дБ}.$$

Минимально реализуемое адаптоаттенюатором ослабление (в предельном случае $R_2/R_1 \approx 0$) так связано с коэффициентом трансформации:

$$q = 20 \lg \left[\sqrt{n^2} + \sqrt{n^2 - 1} \right], \text{ дБ}.$$

В случае строгого отображения (1) после определения параметров m_4, m_5, m_6 по (2) при $m_1 = -1, m_2 = 0, m_3 = 1$ дальнейшие вычисления проводим аналогично структуре рис. 3,б.

Моделирование более сложных топологий ПЭ проведено в отечественном программном комплексе Elcut.

Во всех ПЭ при достаточно сильном взаимодействии входного и выходного контактов ($b/a > 3$, $e/a \approx 1$ для рис.1,а,б; $\rho_1 = \rho_2$ для ПЭ рис.1,д,е; $f/2a \approx 0, e/b \approx 0$ для ПЭ рис.1,в,г) коэффициент трансформации близок к 1.

При $b/a < 1/3$ и прежних условиях для ПЭ их входное и выходное сопротивления с достаточной точностью определяются из соотношений

$$R_{\text{вх}} = \rho K' [\sin(\pi d_1/2b)] / K [\sin(\pi d_1/2b)], \quad R_{\text{вых}} = \rho K' [\sin(\pi d_2/2b)] / K [\sin(\pi d_1/2b)].$$

При этом коэффициент трансформации не превосходит 2-2,5.

При отсутствии взаимодействия входного и выходного контактов коэффициент трансформации определяется степенью несимметрии топологии ПЭ.

Использование кусочно-однородных структур рис.1,д,ж и оптимизация профиля входного и выходного контактов рис.1,е,з при $\rho_1 > \rho_2$ позволяет при малых габаритах расширить диапазон коэффициентов трансформации до 8-10 и в разы уменьшить величины градиентов концентрации и плотности мощности, что важно для импульсного режима работы. Близкие результаты можно получить введением в однородную пленку выреза или области высокой проводимости (ПЭ рис. 1,в,г)

Важными являются и вопросы подгонки параметров ПЭ после изготовления. Параметры схемы замещения образца ПЭ находят путем последовательного замыкания контактов 1-2, 1-3 и 2-3, измерения получающихся сопротивлений R_{12} , R_{13} , R_{23} и вычисляются следующим образом:

$$1/R_1 = 0,5(1/R_{12} + 1/R_{23} - 1/R_{13}), \quad 1/R_2 = 0,5(1/R_{23} + 1/R_{13} - 1/R_{12}), \quad 1/R_3 = 0,5(1/R_{13} + 1/R_{12} - 1/R_{23}).$$

Заключение

Предложены перспективные топологии и методики расчета ПЭ адаптивных аттенуаторов на основе однородных и кусочно-однородных резистивных структур малых габаритов для реализации широкого диапазона ослаблений с заданными значениями входного и выходного сопротивлений для ГИС ВЧ и СВЧ диапазонов, позволяющие в разы уменьшить значения градиентов потенциалов и плотности мощности вблизи входного и выходного контактов, что важно для импульсного режима работы.

Библиографический список

1. Белов Л. Аттенуаторы СВЧ-сигналов. Обзор // Электроника: Наука, Технология, Бизнес. 2006, № 2, с. 32 – 38.
2. Кочемасов В., Белов Л. Фиксированные аттенуаторы - производители и характеристики. Ч. 1 // Электроника НТБ, 2017, № 6, с. 108 - 116.
3. Садков В.Д. Тонкопленочные звенья затухания с трансформацией сопротивлений // Известия вузов СССР – Радиоэлектроника, 1978, т. XXI, № 7, с. 120 – 122.
4. Арамата, Теркадо. К расчету сопротивления одного класса двумерных неоднородных проводников. ТИИЭР, 1969, т.57, № 3, с. 180-181.
5. Ямпурин Н.П., Широков В.Д., Садков В.Д., Потехин В.А. Современные проблемы радиоэлектроники с позиций теории конформных отображений. – Арзамас, АГПИ, 2014. – 209 с.
6. Мартюшов К.И., Зайцев Ю.В., Тихонов А.И. Методы расчета резисторов. – М.: Энергия, 1971. – 208 с.
7. Крылов В.И. Приближенное вычисление интегралов. – М.: Наука, 1967. -500 с.
8. Татур Г.А. Основы теории электрических цепей (справочное пособие): Учебное пособие. - М.: Высшая школа, 1980. – 271 с.

THIN-FILM ABSORBING ELEMENTS OF ADAPTOATTENUATORS FOR GIS

Nizhny Novgorod state technical University n.a.R.E. Alekseeva

Promising topologies and methods for calculating the absorbing elements of adaptoattenuators – film absorbing elements with different values of input and output resistances based on homogeneous and piecewise homogeneous resistive structures for implementing a wide range of attenuations with specified values of input and output resistances for hybrid integrated circuits are proposed.

Topologies of absorbing elements based on a homogeneous resistive film with a conductive or non-conductive region are considered, which allow providing the required range of attenuations and transformation coefficients in small dimensions.

The investigated topology is based on piecewise-uniform resistive film, including the optimal profile of input and output contacts, allowing not only to reduce the sizes of the absorbing elements, but in 2-3 times to reduce the maximum value of the potential gradient and power density, which is especially important for pulsed operation of adapteractivator.

Keywords: Film absorbing element, adaptationfor, gradient, potential, power density, an optimum profile.

2LЭКСПРЕСС МЕТОД ИСКЛЮЧЕНИЯ ВЛИЯНИЯ ОКРУЖАЮЩИХ ЦЕПЕЙ В НЕСТАНДАРТНЫХ НАПРАВЛЯЮЩИХ СИСТЕМАХ

Нижегородский государственный технический университет им. Р.Е. Алексеева

В работе рассмотрен 2LЭкспресс метод измерения S-параметров устройств, являющийся развитием LRT, LR и Экспрессметодов. Предлагаемый метод решает проблему в определении средней линии, являющейся «тонким» местом этих методов.

Ключевые слова: коаксиально-полосковый переход, полосковые линии передачи

Контроль параметров устройств относительно его физических границ сводится к исключению влияния окружающих его электрических цепей. Эта процедура называется калибровкой. Существует много известных алгоритмов калибровки. Например, калибровка OSM позволяет калиброваться в широком частотном диапазоне, но требует наличия меры согласованной нагрузки. Для нестандартных линий передачи такую меру изготовить или крайне сложно, или практически невозможно. Еще один широко известный метод TRN не требует меры согласованной нагрузки, но является узкополосным. В работе [1] был разработан оригинальный LRT метод калибровки, который не требует наличия дорогой и сложной в изготовлении меры согласованной нагрузки и работает в широкой полосе частот, затем этот метод получил развитие в работе [2] как LR метод калибровки, этот метод позволяет выполнить калибровку всего на двух мерах, а в работе [3] был предложен Экспресс метод, в которой для калибровки оказалось достаточно всего одной меры электрически длинной линии. Все новые методы имеют общий методический недостаток – определение коэффициентов отражения со стороны коаксиальных разъемов L-соединений переходов как средних линий комплексной функции частоты электромагнитного поля. Средние линии, полученные разными алгоритмами, имеют достаточно сильные погрешности. На рис. 1 показано, насколько сильно отличаются полученные средние линии. Синий график получен методом минимакса, зеленый график – методом МНК.

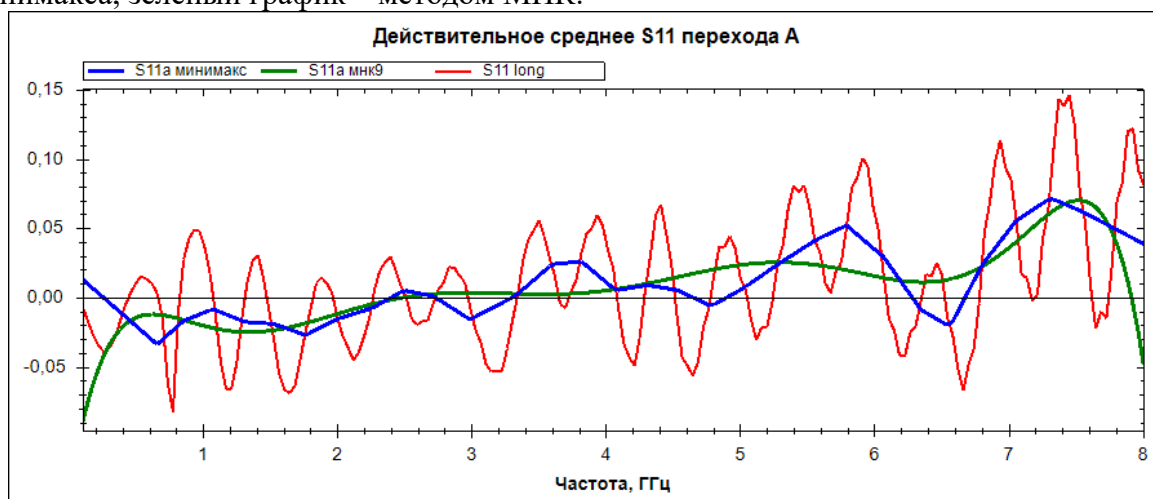


Рис. 1. Сравнение средних линий, полученных разными методами

Видно, что средняя линия, полученная методом минимакса, имеет ломаный характер, а средняя линия, полученная алгоритмом МНК, не учитывает выбросы.

В настоящей работе предлагается новый метод исключения КНП как развитие LRT, LR и Экспресс методов. Метод основан на получении базовых точек, в которых и выполняется начальная калибровка.

В данной работе предлагается использовать новый 2LЭкспресс метод контроля параметров, позволяющий получать среднюю линию аналитическим путем. То есть метод, сохраняя все достоинства LRT, LRи Экспресс методов, устраняет их основной недостаток – неточность определения средней линии. Метод является универсальным, позволяет работать как в стандартных, так и в нестандартных направляющих системах.

Для работы с методом необходимо провести три измерения:

- с отрезком электрически длинной линии, получить S-параметры S_{ik}^L ;
- с отрезком электрически короткой линии, получить S-параметры S_{ik}^T ;
- с отрезками линий в режиме короткого замыкания или холостого хода, получить S-параметры S_{ik}^R .

Определить коэффициенты отражения со стороны коаксиальных разъемов S_{11}^a и S_{11}^b в базовых точках, где фаза S_{21}^L/S_{21}^T принимает значение $\pm 90^\circ$ из соотношений:

$$S_{11}^a = \frac{S_{11}^T S_{12}^L - S_{11}^L S_{12}^T T}{S_{12}^L - S_{12}^T T}, S_{11}^b = \frac{S_{22}^T S_{21}^L - S_{22}^L S_{21}^T T}{S_{21}^L - S_{21}^T T} \quad (1)$$

Здесь: $T = \exp(-i\gamma l)$, $l = L - l_1$ l_1 – длина короткой линии,

Измерения проводились на векторном анализаторе цепей P4226 «Панорама» в диапазоне частот 0.01 – 26 ГГц. Для подключения линий к анализаторам цепей были использованы два коаксиально-полосковые перехода 292-04А-5 с коаксиальным соединителем типа SMA (производитель Southwest Microwave Inc.). Длинная линия имела длину 120 мм, короткая – 30 мм. Обе линии изготовлены из фольгированного материала RO4003С, поперечные размеры выбраны с помощью калькулятора AppCAD-[CPW] так, чтобы волновое сопротивление имело величину 50 Ом на частоте 20 ГГц.

По отношению коэффициентов передачи длинной и короткой линии был определен массив базовых точек, показанный на рис. 2.

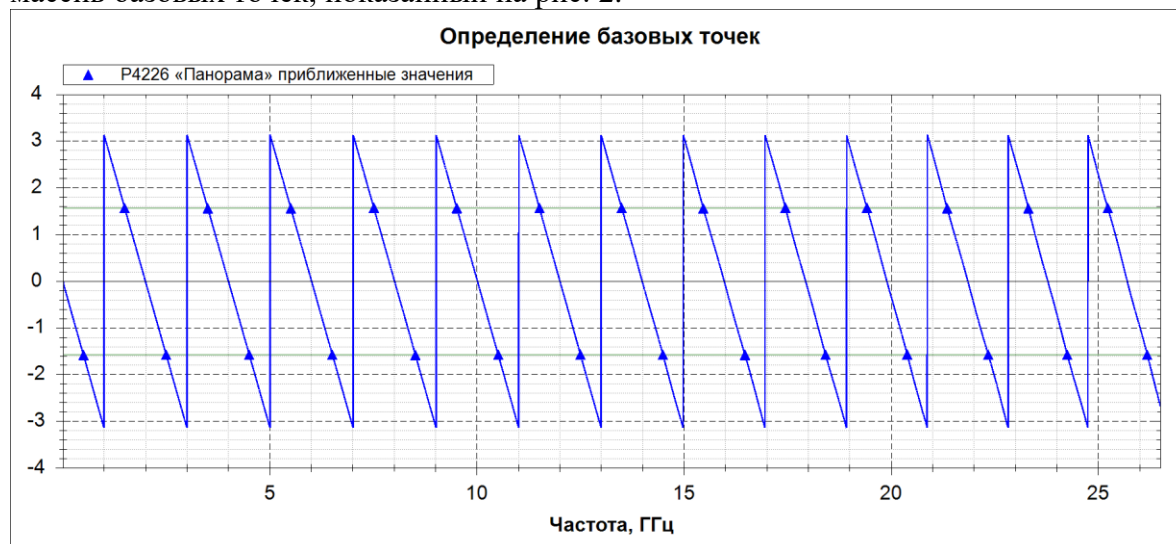


Рис. 2. График фазы отношения S_{21}^L/S_{21}^T с базовыми точками

В базовых точках дифференциально-фазовым методом были определены эффективная диэлектрическая проницаемость и коэффициент затухания полосковой линии из отношения:

$$\alpha_n = -\frac{\text{Ln}(T_n)}{\Delta L}, \varepsilon_{eff n} = \left(\frac{2n-1}{4f_n \Delta L} v_0 \right)^2, \Delta L = 90 \text{ мм} \quad (2)$$

Затем из отношения (1) были рассчитаны коэффициенты отражения коаксиально-полосковых переходов со стороны коаксиального разъема. После этого была выполнена

кусочную интерполяцию всего диапазона. В работе [5] проведено сравнение трех методов кусочной интерполяции: линейная, кубическая и сплайнами Акимы. Полученные результаты оценивались как по точности, так и по времени получения результата. Показано, что для разреженных точек и данных, имеющих выбросы, лучше всего воспользоваться сплайнами Акимы, так как они устойчивы к выбросам и на разреженных данных не дают «ложных выбросов», как, например, наиболее популярные кубические сплайны.

После получения значения коэффициентов отражения со стороны коаксиальных разъемов S_{11}^a и S_{11}^b во всем диапазоне для завершения процедуры исключения влияния переходов на результаты измерения используются следующие отношения:

$$\begin{aligned} s_{22}^a(f) &= K \frac{s_{22}(f) - s_{11}^b(f)}{s_{12}(f)}, \quad s_{22}^b(f) = \frac{s_{11}(f) - s_{11}^a(f)}{s_{21}(f)K}, \\ s_{21}^a(f)s_{12}^a(f) &= s_{21}(f)K(1 - s_{22}^a(f)s_{22}^b(f)), \\ s_{21}^b(f)s_{12}^b(f) &= s_{12}(f)(1 - s_{22}^a(f)s_{22}^b(f))/K, \\ s_{21}^a(f)s_{12}^b(f) &= s_{21}(f)(1 - s_{22}^a(f)s_{22}^b(f)), \\ s_{21}^b(f)s_{12}^a(f) &= s_{12}(f)(1 - s_{22}^a(f)s_{22}^b(f)). \end{aligned} \quad (3)$$

Коэффициент $K = s_{21}^a(f)/s_{21}^b(f)$, в отличие от классического Экспресс метода, получается из отношений:

$$\begin{aligned} K(f) &= \frac{\sqrt{R_a} (1 + R_b + iX_b)}{\sqrt{R_b} (1 + R_a + iX_a)}, \\ X_{a,b}(f) &= \frac{2 \operatorname{Im}(s_{11}^{a,b})}{1 - 2 \operatorname{Re}(s_{11}^{a,b}) + |s_{11}^{a,b}|^2}, \quad R_{a,b}(f) = \frac{1 - |s_{11}^{a,b}|^2}{1 - 2 \operatorname{Re}(s_{11}^{a,b}) + |s_{11}^{a,b}|^2}. \end{aligned} \quad (4)$$

После этого выполняется процедура исключения S-параметров цепей А и В из результатов измерений S-параметров короткой линии, определяется точное значение коэффициента передачи T отрезка линии длиной $\Delta L = 90 \text{ мм}$, затем уточняется массив частотных точек (рисунок 3) и уже на точном массиве с помощью отношения (2) получают окончательные оценки искомых величин $\alpha(f)$ и $\varepsilon_{\text{eff}}(f)$.

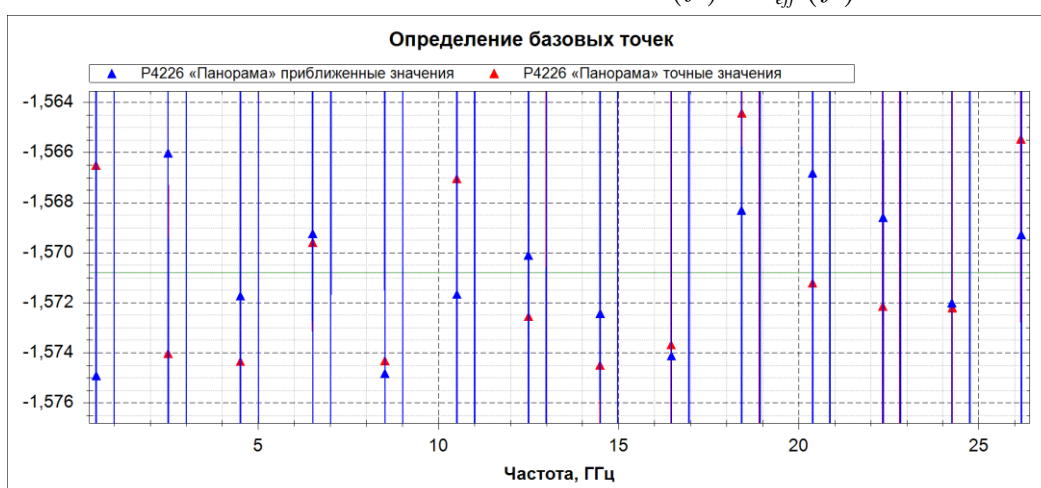


Рис. 3. Сравнение приближенного и точного массивов базовых точек

Был проведен анализ отклонения значения фазы отношения коэффициентов передачи длинной и короткой полосковых линий от величины $\pm\pi/2$ и значения фазы коэффициента передачи T отрезка линии длиной $\Delta L = 90 \text{ мм}$. Среднее отклонение фазы

отношения T коэффициентов передачи S_{21} отрезков длиной и короткой линий для ВАЦ Р4226 «Панорама» составило 0,07475 или 0,1762%. Для фазы коэффициента передачи T отрезка линии длиной $\Delta L = 90 \text{ мм}$ среднее отклонение оказалось практически таким же 0,07470 или 0,1761%.

После уточнения массива базовых точек и значений коэффициентов затухания и эффективной диэлектрической проницаемости были окончательно рассчитаны остальные параметры цепей А и В, а именно коэффициенты отражения $s_{22}^a(f)$ и $s_{22}^b(f)$ со стороны их полосковых портов, произведения прямого и обратного коэффициентов передачи $s_{21}^a(f)s_{12}^a(f)$ и $s_{21}^b(f)s_{12}^b(f)$, а также произведения коэффициентов передачи $s_{21}^a(f)s_{12}^b(f)$ и $s_{21}^b(f)s_{12}^a(f)$ по отношениям (3-4).

Для оценки получившихся значений те же параметры были рассчитаны 2LR и 2LOS методами.

2LR метод является вариантом LR метода [3], где для определения коэффициентов отражения со стороны коаксиальных разъемов S_{11}^a и S_{11}^b используется метод 2L.

2LOS метод – вариант OSM метода, где используются только короткозамкнутая мера и мера холостого хода, предварительно аттестованные с помощью 2L Экспресс метода. Для аттестации короткозамкнутой нагрузки из измерения меры короткого замыкания были вычтены окружающие цепи А и В, S-параметры которых были получены с помощью 2L Экспресс метода по соотношениям (3) и (4). Затем был определен коэффициент отражения от короткозамкнутой нагрузки по формуле:

$$\Gamma_{sh} = \frac{S_{11}^{sh} - S_{11}^a}{S_{11}^{sh} S_{22}^a - \det S^a}. \quad (5)$$

Аналогично был получен коэффициент отражения от нагрузки холостого хода из соотношения:

$$\Gamma_{op} = \frac{S_{11}^{op} - S_{11}^a}{S_{11}^{op} S_{22}^a - \det S^a}. \quad (6)$$

Затем остальные параметры цепей А и В рассчитываются из соотношений

$$\begin{aligned} S_{22}^a &= \frac{(S_{11}^a - S_{11}^{op})\Gamma_{sh} - (S_{11}^a - S_{11}^{sh})\Gamma_{op}}{(S_{11}^{sh} - S_{11}^{op})\Gamma_{sh}\Gamma_{op}}, \\ S_{22}^b &= \frac{(S_{11}^b - S_{22}^{op})\Gamma_{sh} - (S_{11}^b - S_{22}^{sh})\Gamma_{op}}{(S_{22}^{sh} - S_{22}^{op})\Gamma_{sh}\Gamma_{op}}, \\ \det Sa &= \frac{(S_{11}^a - S_{11}^{op})S_{11}^{sh}\Gamma_{sh} - (S_{11}^a - S_{11}^{sh})S_{11}^{op}\Gamma_{op}}{(S_{11}^{sh} - S_{11}^{op})\Gamma_{sh}\Gamma_{op}}, \\ \det Sb &= \frac{(S_{11}^b - S_{22}^{op})S_{22}^{sh}\Gamma_{sh} - (S_{11}^b - S_{22}^{sh})S_{22}^{op}\Gamma_{op}}{(S_{22}^{sh} - S_{22}^{op})\Gamma_{sh}\Gamma_{op}}, \\ s_{21}^a s_{12}^b &= S_{21}^L (1 - S_{22}^a S_{22}^b). \end{aligned} \quad (7)$$

На рисунках 4-6 показано сравнение коэффициента отражения цепи А со стороны полосковой линии, рассчитанного разными методами – 2L Экспресс, 2LR и 2LOS методами. Хорошо видно, что действительное и мнимое коэффициента отражения, полученного 2L Экспресс методом, имеет более гладкий характер, чем тот же коэффициент отражения, рассчитанный 2LR методом. Это можно объяснить тем, что при расчетах 2LR методом необходимо проводить еще одно измерение мерой короткого замыкания или холостого хода, и здесь проявляется неоднородность соединителей, на

которую указывал Дансмор Джоэль в работе [6]. А вот модуль коэффициента отражения полностью совпадает, что говорит о высокой точности обоих методов.

Результаты, полученные 2LOS и 2L Экспресс методами практически полностью совпадают. Это объясняется тем, что коэффициенты отражения нагрузок короткого замыкания и холостого хода были получены по измерениям с помощью 2L Экспресс метода.

Предложенный в данной работе 2L метод определения средней линии позволяет устранить методический недостаток LRI Экспресс методов, а изменения, связанные с определением отношения коэффициентов передачи, позволяют получить менее зашумленные результаты.

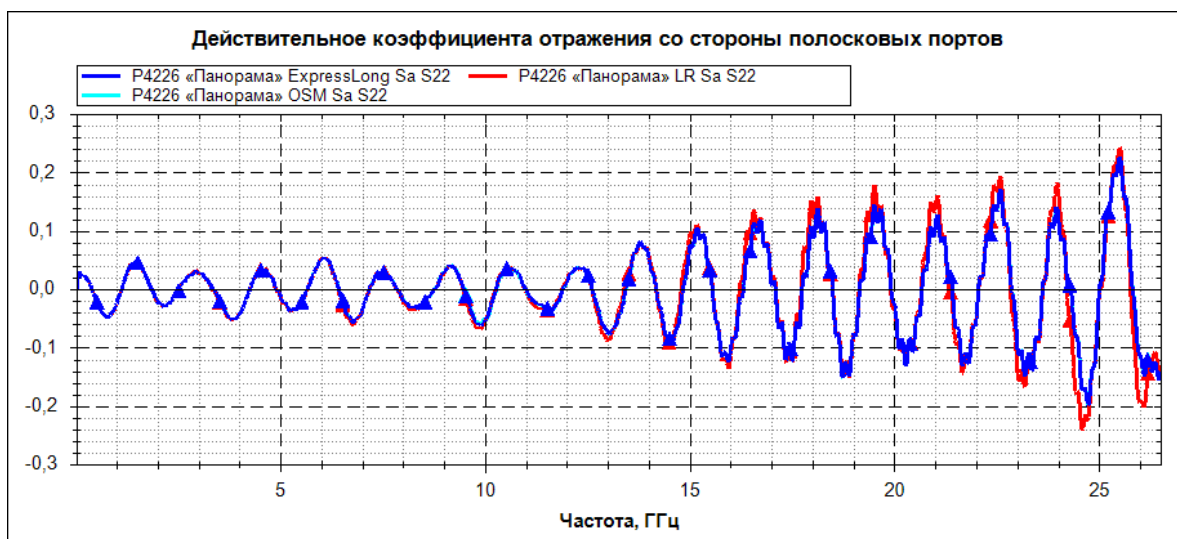


Рис. 4. Сравнение действительного коэффициента отражения, полученного разными методами

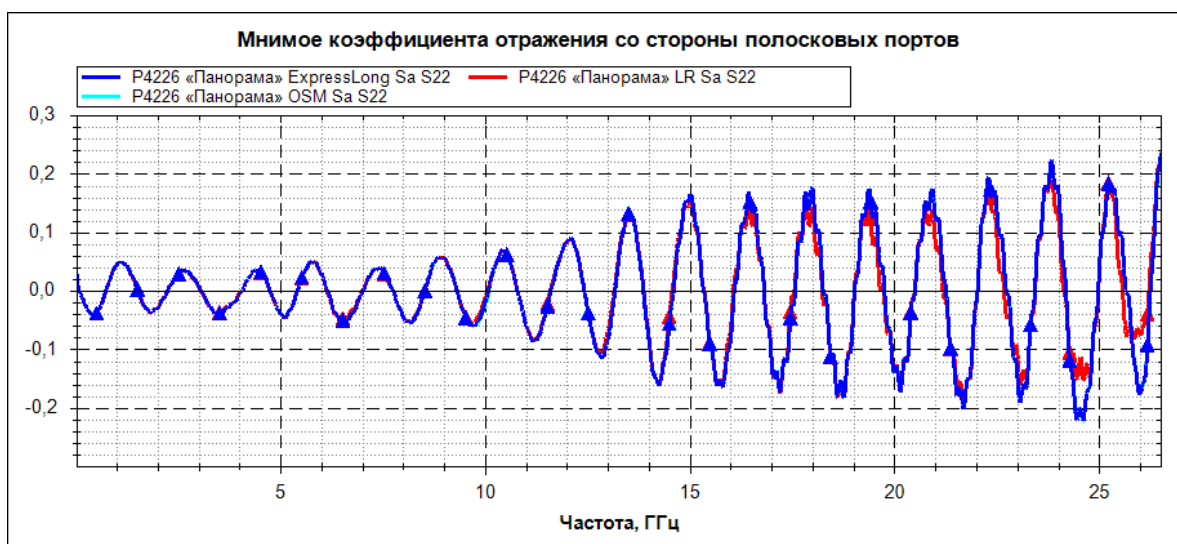


Рис. 5. Сравнение мнимого коэффициента отражения, полученного разными методами

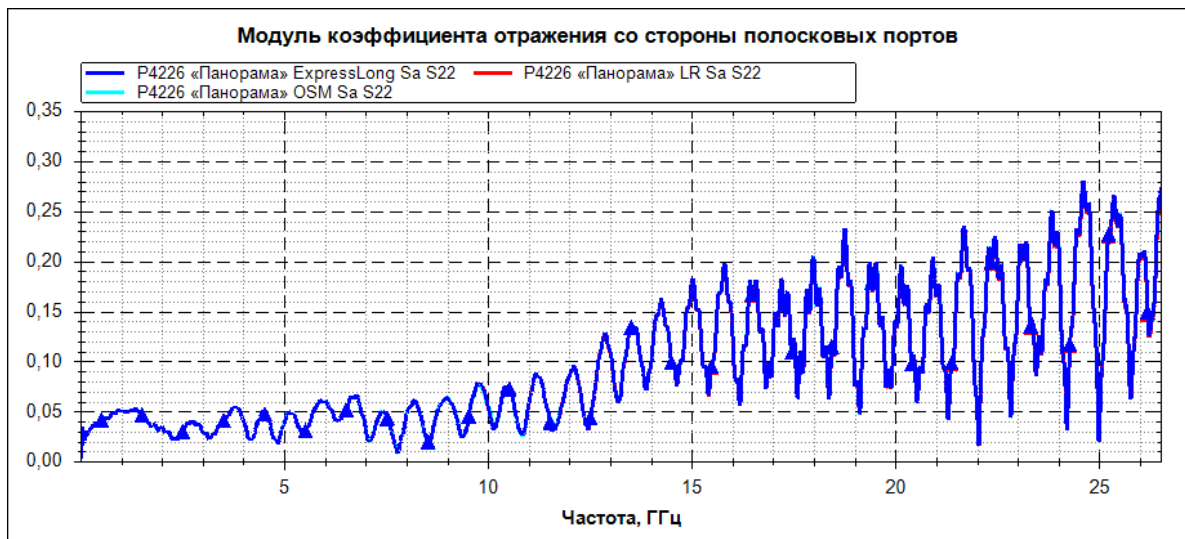


Рис. 6. Сравнение модуля коэффициента отражения, полученного разными методами

Библиографический список

1. Лавричев О.В. LRT-метод определения параметров объектов в нестандартных направляющих системах / О.В. Лавричев, С.М, Никулин // Датчики и системы. – М.: Изд-во Сенсидат-Плюс 2017, №8-9 с. 39-44.
2. Евсеев В.И. Контроль параметров объектов в нестандартных направляющих системах и в свободном пространстве / В.И. Евсеев [и др.] //UralRadioEngineeringJournal, Уральский радиотехнический журнал. 2019. Vol. 3, No 1. p. 53–71.
3. Евсеев В.И. LR и экспресс-метод контроля параметров электронных компонентов в полосковых линиях передачи / В.И. Евсеев, С.М. Никулин // Электроника и микροэлектроника СВЧ. 2020, с. 18 -22 ISBN978-5-7629-2456-6.
4. Das N.K. Two Methods for the Measurement of Substrate Dielectric Constant / N.K. Das, S.M. Voda and D.M. Pozar// IEEE Transactions on Microwave Theory and Techniques. Vol. 35. No 7. July 1987.
5. Круковец А.С. Разработка метода интерполяции значений номограммы / А.С. Круковец, Г.А. Горелкин // Современные научные исследования и инновации. 2015. № 5. Ч. 2 [Электронный ресурс]. URL: <http://web.snauka.ru/issues/2015/05/53846> (дата обращения: 02.02.2020).
6. Дансмор Джоэль П. Настольная книга инженера. Измерение параметров СВЧ устройств с использованием передовых методик векторного анализа цепей /Джоэль П.Дансмор// Москва: ТЕХНОСФЕРА. 2018. – 736 с.

Е.А. Lupanova

2L EXPRESS CIRCUIT REDUCTION METHOD IN NON-STANDARD GUIDE SYSTEMS

Nizhny Novgorod state technical university n.a. R.E. Alekseev

The paper considers the 2L Express method for measuring the S-parameters of devices, which is a development of the LRT, LR and Express methods. The proposed method solves the problem of determining the midline, which is the "thin" place of these methods.

Keywords: coaxial strip junction, strip transmission lines.

ОПРЕДЕЛЕНИЕ ЭФФЕКТИВНОЙ ДИЭЛЕКТРИЧЕСКОЙ ПРОНИЦАЕМОСТИ И КОЭФФИЦИЕНТА ЗАТУХАНИЯ В ПОЛОСКОВЫХ ЛИНИЯХ ПЕРЕДАЧИ

Нижегородский государственный технический университет им. Р.Е. Алексеева

В работе предложен новый метод определения собственных параметров полосковых линий передачи на основе 2L Экспресс метода.

Ключевые слова: коаксиально-полосковый переход, полосковые линии передачи, эффективная диэлектрическая проницаемость, коэффициент затухания

В литературе известны различные функциональные соотношения для волнового сопротивления и эффективной диэлектрической проницаемости $\varepsilon_{eff}(f)$ в полосковых линиях передачи [1]. Причем частотную зависимость $\varepsilon_{eff}(f)$ удастся определить дифференциально-фазовым методом по результатам измерений в коаксиальном канале коэффициентов передачи длинной и короткой полосковых линий [2]. Однако, даже при использовании одних и тех же коаксиально-полосковых переходов, подключаемых к одним и тем же портам анализатора цепей, возникают неустраняемые погрешности, вызванные методической погрешностью, а также повторяемостью коаксиальных и коаксиально-полосковых разъемных соединений [3]. Кроме того, параметры полосковых линий определяется их конструкцией, геометрическими размерами поперечного сечения, относительной диэлектрической проницаемостью изоляционного основания и другими факторами, которые не поддаются точному оцениванию и аналитическому моделированию. По этой причине поиск корректных методов определения собственных параметров полосковых линий средствами натурального эксперимента продолжается. Искомые величины в скрытой форме присутствуют в S-параметрах соединения двух коаксиально-полосковых переходов электрически длинной полосковой линией передачи. В настоящей работе предлагается способ извлечения этой информации с целью решения поставленной задачи.

Ключевым фактором, определяющим S-параметры соединения, являются области контакта центрального проводника коаксиальной линии с полосковым проводником. Ввиду достаточно малых продольных размеров этих областей, их можно представить на электрической схеме, показанной на рисунке 1, реактивными импедансами $X_a(f)$ и $X_b(f)$, моделирующими трансформацию электромагнитных полей коаксиальной и полосковой линий передачи. Коаксиальные области конструкции переходов моделируются идеальными линиями передачи длиной L_a и L_b с воздушным диэлектриком, а полосковая линия длиной L имеет частотно-зависимые волновое сопротивление $Z(f)$ и коэффициент затухания $\alpha(f)$. Дисперсия фазовой скорости $v(f) = v_0 / \sqrt{\varepsilon_{eff}(f)}$ электромагнитной волны в полосковой линии определяется относительной эффективной диэлектрической проницаемостью $\varepsilon_{eff}(f)$, v_0 - скорость волны в свободном пространстве.

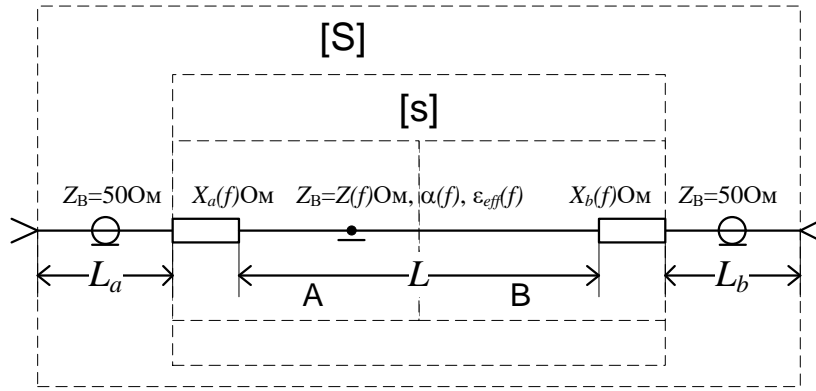


Рис. 1. Схемная модель соединения коаксиально-полосковых переходов с отрезком полосковой линии

$[S]$ - S-параметры соединения коаксиально-полосковых переходов с полосковой линией,
 $[s]$ - S-параметры соединения без отрезков коаксиальных линий длиной L_a и L_b , А, В– двух
 портовые цепи с коаксиальными входами и полосковыми выходами,
 $X_a(f)$ и $X_b(f)$ - реактивные импедансы

Участок схемы, имеющий матрицу рассеяния $[s]$, представляет собой каскадное соединение цепей А и В, включенных навстречу друг другу. Коэффициенты отражения $s_{11}^a(f)$ и $s_{11}^b(f)$ со стороны их коаксиальных портов в случае электрически длинной полосковой линии являются средними линиями квазипериодических комплексных функций частоты $s_{11}(f)$ и $s_{22}(f)$ [4]. Остальные параметры цепей А и В: - коэффициенты отражения $s_{22}^a(f)$ и $s_{22}^b(f)$ со стороны их полосковых портов, произведения прямого и обратного коэффициентов передачи $s_{21}^a(f)s_{12}^a(f)$ и $s_{21}^b(f)s_{12}^b(f)$, а также произведения коэффициентов передачи $s_{21}^a(f)s_{12}^b(f)$ и $s_{21}^b(f)s_{12}^a(f)$ цепей А и В, определяются Экспресс-методом из соотношений [5]:

$$\begin{aligned}
 s_{22}^a(f) &= K \frac{s_{22}(f) - s_{11}^b(f)}{s_{12}(f)}, \quad s_{22}^b(f) = \frac{s_{11}(f) - s_{11}^a(f)}{s_{21}(f)K}, \\
 s_{21}^a(f)s_{12}^a(f) &= s_{21}(f)K(1 - s_{22}^a(f)s_{22}^b(f)), \\
 s_{21}^b(f)s_{12}^b(f) &= s_{12}(f)(1 - s_{22}^a(f)s_{22}^b(f))/K, \\
 s_{21}^a(f)s_{12}^b(f) &= s_{21}(f)(1 - s_{22}^a(f)s_{22}^b(f)), \\
 s_{21}^b(f)s_{12}^a(f) &= s_{12}(f)(1 - s_{22}^a(f)s_{22}^b(f)).
 \end{aligned} \tag{1}$$

Коэффициент $K = s_{21}^a(f)/s_{21}^b(f)$, равный отношению коэффициентов передачи цепей А и В, и реактивные импедансы $X_a(f)$ и $X_b(f)$ определяют из соотношений:

$$\begin{aligned}
 K(f) &= \frac{\sqrt{R_a} (1 + R_b + iX_b)}{\sqrt{R_b} (1 + R_a + iX_a)}, \\
 X_{a,b}(f) &= \frac{2\text{Im}(s_{11}^{a,b})}{1 - 2\text{Re}(s_{11}^{a,b}) + |s_{11}^{a,b}|^2}, \quad R_{a,b}(f) = \frac{1 - |s_{11}^{a,b}|^2}{1 - 2\text{Re}(s_{11}^{a,b}) + |s_{11}^{a,b}|^2}.
 \end{aligned} \tag{2}$$

Отношение (2) является принципиальным отличием от классического Экспресс метода, где коэффициент K принимается равным 1.

Собственные параметры $\alpha(f)$ и $\varepsilon_{eff}(f)$ экранированной полосковой линии определены по S-параметрам, измеренным в коаксиальном канале анализатором цепей PNA N5222A в диапазоне частот 0.01 – 26.5 ГГц. В качестве объектов измерений, как и в дифференциально-фазовом методе [2], использованы две линии: - длинная, короткая. Для подключения линий к анализатору цепей использованы коаксиально-полосковые переходы 292-04А-5 с коаксиальным соединителем типа SMA (производитель Southwest Microwave Inc.). Длинная линия имела длину 120 мм, короткая – 30 мм. Обе линии изготовлены из фольгированного материала RO4003С, поперечные размеры выбраны с помощью калькулятора AppCAD-[CPW] так, чтобы волновое сопротивление имело величину 50 Ом на частоте 20 ГГц.

На первом этапе были определены коэффициенты отражения со стороны коаксиальных разъемов S_{11}^a и S_{11}^b в базовых точках, где фаза S_{21}^L/S_{21}^T принимает значение $\pm 90^\circ$ из соотношений:

$$S_{11}^a = \frac{S_{11}^T S_{12}^L - S_{11}^L S_{12}^T T}{S_{12}^L - S_{12}^T T}, S_{11}^b = \frac{S_{22}^T S_{21}^L - S_{22}^L S_{21}^T T}{S_{21}^L - S_{21}^T T} \quad (1)$$

Здесь: $T = \exp(-i\gamma l)$, $l = L - l_t$, l_t – длина короткой линии,

Полученный массив базовых точек показан на рисунке 2.

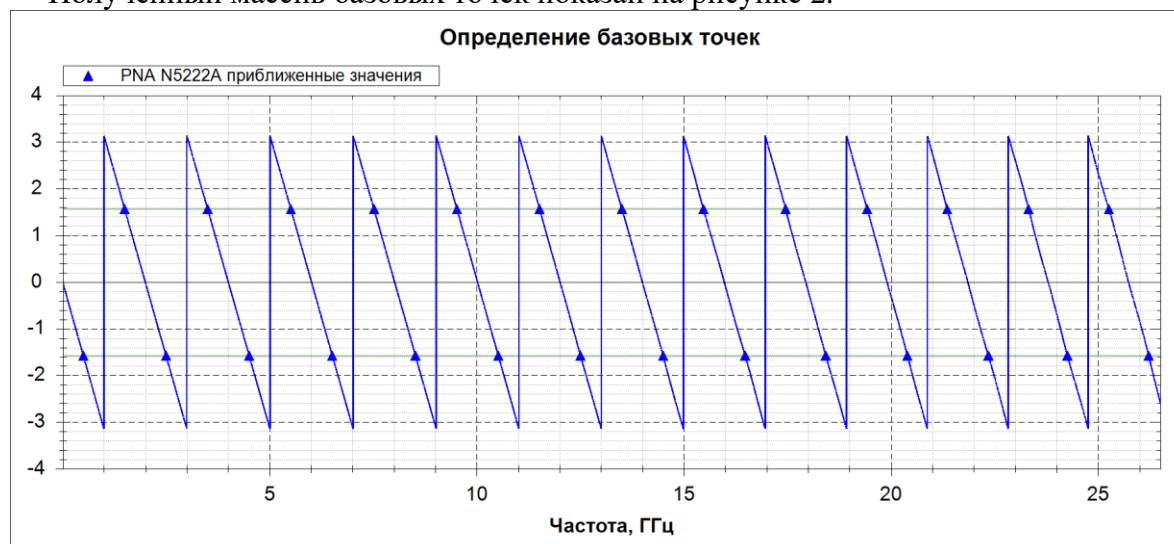


Рис. 2. График фазы отношения S_{21}^L/S_{21}^T с базовыми точками

В базовых точках дифференциально-фазовым методом были определены эффективная диэлектрическая проницаемость и коэффициент затухания полосковой линии из отношения:

$$\alpha_n = -\frac{\text{Ln}(T_n)}{\Delta L}, \varepsilon_{eff n} = \left(\frac{2n-1}{4f_n \Delta L} v_0 \right)^2, \Delta L = 90 \text{ мм} \quad (2)$$

Затем из отношения (1) были рассчитаны коэффициенты отражения коаксиально-полосковых переходов со стороны коаксиального разъема. После этого была выполнена кусочная интерполяция всего диапазона. В работе [6] проведено сравнение трех методов кусочной интерполяции: линейная, кубическая и сплайнами Акимы. Полученные результаты оценивались как по точности, так и по времени получения результата. Показано, что для разреженных точек и данных, имеющих выбросы, лучше всего воспользоваться сплайнами Акимы, так как они устойчивы к выбросам и на разреженных данных не дают «ложных выбросов», как, например, наиболее популярные кубические сплайны.

После этого выполняется процедура исключения S-параметров цепей А и В из результатов измерений S-параметров короткой линии, определяется точное значение коэффициента передачи T отрезка линии длиной $\Delta L = 90 \text{ мм}$, затем уточняется массив частотных точек и уже на точном массиве с помощью отношения (2) получаются окончательные оценки искомых величин $\alpha(f)$ и $\varepsilon_{\text{eff}}(f)$.

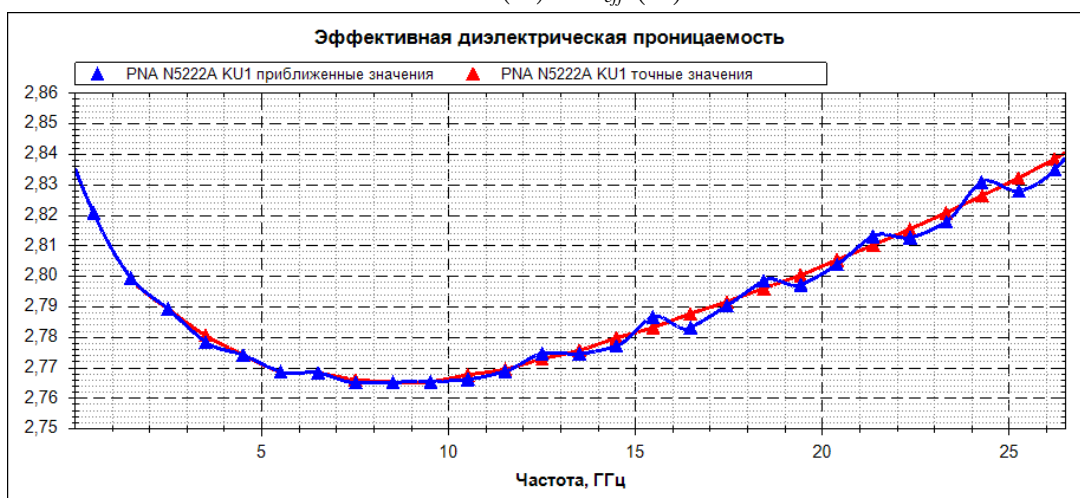


Рис. 3. Сравнение приближенных и точных значений эффективной диэлектрической проницаемости

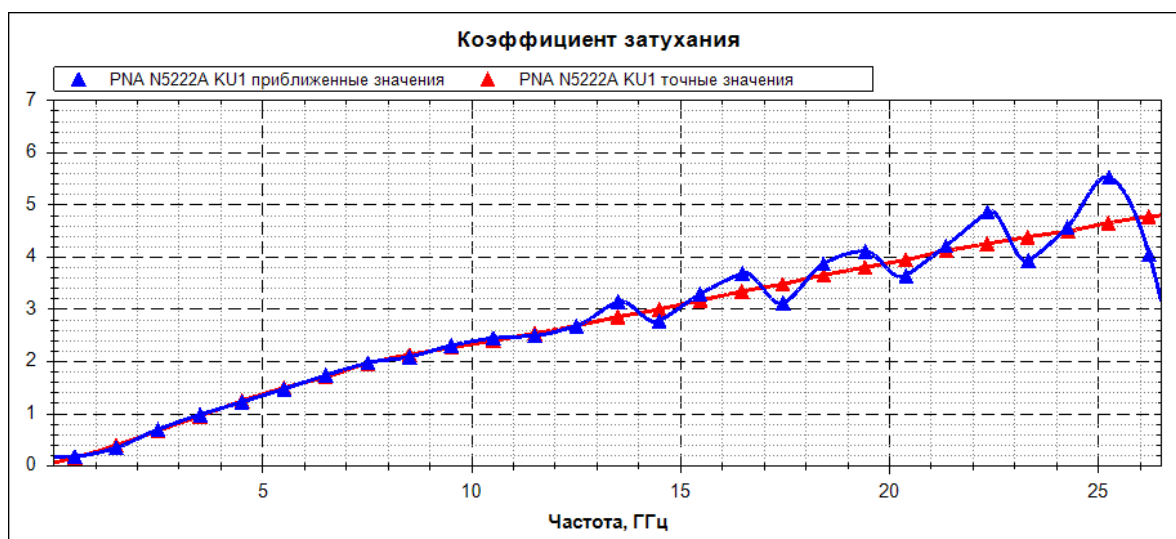


Рис. 4. Сравнение приближенных и точных значений коэффициента затухания

На рисунках хорошо видно, что график приближенных значений как эффективной диэлектрической проницаемости, так и коэффициента затухания, с увеличением частоты приобретает ярко выраженный колебательный характер.

Предложенный в данной работе метод определения эффективной диэлектрической проницаемости и коэффициента затухания полосковых линий передачи позволяет получить значительно лучшие результаты, чем дифференциально-фазовый метод благодаря процедуре предварительно калибровки, выполненной с помощью 2L Экспресс метода.

Библиографический список

1. Wadell Brian C. Transmission line design handbook / Brian C. Wadell, p. 513. Includes bibliographical references and index. ISBN 0-89006-436-9.
2. Das N.K. Two Methods for the Measurement of Substrate Dielectric Constant / N.K. Das, S.M. Voda and D.M. Pozar // IEEE Transactions on Microwave Theory and Techniques. Vol. 35. No 7. July 1987.
3. Дансмор Джоэль П. Настольная книга инженера. Измерение параметров СВЧ устройств с использованием передовых методик векторного анализа цепей / Джоэль П. Дансмор // Москва: ТЕХНОСФЕРА. 2018. – 736 с.
4. Лавричев О.В. LRT-метод определения параметров объектов в нестандартных направляющих системах / О.В. Лавричев, С.М. Никулин // Датчики и системы. – М.: Изд-во Сенсидат-Плюс 2017, №8-9 с. 39-44.
5. Евсеев В.И. LR и экспресс-метод контроля параметров электронных компонентов в полосковых линиях передачи / В.И. Евсеев, С.М. Никулин // Электроника и микроэлектроника СВЧ. 2020, с. 18 -22 ISBN978-5-7629-2456-6.
6. Круковец А.С. Разработка метода интерполяции значений номограммы / А.С. Круковец, Г.А. Горелкин // Современные научные исследования и инновации. 2015. № 5. Ч. 2 [Электронный ресурс]. URL: <http://web.snauka.ru/issues/2015/05/53846> (дата обращения: 02.02.2020).

Е.А. Lupanova

DETERMINATION OF EFFECTIVE PERMITTIVITY AND ATTENUATION COEFFICIENT IN STRIP TRANSMISSION LINES

Nizhny Novgorod State Technical University n.a. R.E. Alekseev

The paper proposes a new method for determining the intrinsic parameters of strip transmission lines based on the 2L Express method.

Keywords: coaxial strip junction, strip transmission lines, effective permittivity, attenuation coefficient.

LM МЕТОД КОНТРОЛЯ ПАРАМЕТРОВ ОБЪЕКТОВ В СВОБОДНОМ ПРОСТРАНСТВЕ

Нижегородский государственный технический университет им. Р.Е. Алексеева

Обсуждается LM-метод контроля электромагнитных характеристик материалов, поглощающих покрытий и бистатической эффективной площади рассеяния малоразмерных объектов. Принципиальная особенность предлагаемого решения по сравнению с существующими методами антенных измерений состоит в исключении влияния на получаемые результаты характеристик передающих и приемных антенн, включая окружающие пространства между антеннами и объектом контроля. Кроме того, метод позволяет исключить влияние непрямых трасс распространения электромагнитных волн.

Ключевые слова: свободное пространство, электромагнитные характеристики материалов, бистатическая эффективная площадь рассеяния малоразмерных объектов

Электромагнитные характеристики объектов в открытом пространстве определяют методами дистанционного зондирования при измерении комплексных коэффициентов отражения и передачи с помощью анализаторов цепей. Обработку получаемой измерительной информации выполняют методами, основанными на преобразовании Фурье и оконной фильтрации [1-6]. Такие измерения проводят в широкой полосе частот, что предполагает использование достаточно дорогого оборудования, в частности, широкополосных антенн. Зондирование объектов электромагнитной волной в открытом пространстве выполняют с помощью направленных антенн. Для локализации излучения антенн в заданном телесном угле используют фокусирующие линзы и зеркальные отражатели – коллиматоры [7]. Использование таких фокусирующих средств усложняет и удорожает используемое оборудование антенных измерений.

В данной работе предлагается LM метод контроля электромагнитных характеристик объектов в открытом пространстве на основе LR-метода, позволяющий исключить влияние окружающего пространства и направленных антенн на результаты измерений контролируемых объектов.

Функцию переходов здесь выполняют две направленные антенны, подключенные к векторному анализатору цепей, а функцию нестандартной линии передачи – открытое пространство между антеннами. Для формирования электромагнитной волны с плоским фазовым фронтом и постоянной амплитудой необходимо иметь достаточно большое расстояние между антенной и контролируемым объектом. Расстояние между антеннами удается предельно сократить, как показано на рис. 1, при наклонном падении зондирующей волны на удаленный отражатель: - плоское зеркало или контролируемый объект.

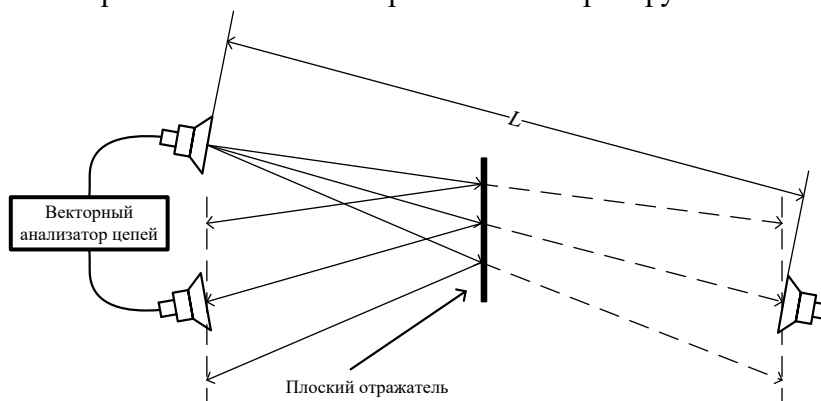


Рис. 1. L-соединение антенн в открытом пространстве

Такое расположение антенн позволяет измерить бистатистический (двухпозиционный) комплексный коэффициент отражения от плоских образцов материалов и поглощающих покрытий и малоразмерных объектов. Антенны соединяют с векторным анализатором цепей короткими коаксиальными кабелями. Оптические оси антенн ориентированы таким образом, чтобы сигнал одной антенны принимался другой антенной за счет отражения от зеркала или контролируемого объекта, как показано на рис. 1.

Каждая антенна с полупространством от антенны до зеркала образует электрическую цепь. Эти электрические цепи, соответственно А и В, оказываются включенными навстречу друг другу.

Расстояние от антенн до зеркала $L/2$ выбирают таким образом, чтобы зеркало, а затем и измеряемый объект (пластина плоского материала ограниченных размеров, пластина с отражающим покрытием или малоразмерный объект) оказались под воздействием квазиплоского электромагнитного излучения в области главных лепестков диаграмм направленности антенн. Тем самым моделируются условия, близкие к зондированию объектов измерения наклонно падающей плоской электромагнитной волной.

Математическое описание L-соединения в открытом пространстве поясняет рис. 2.

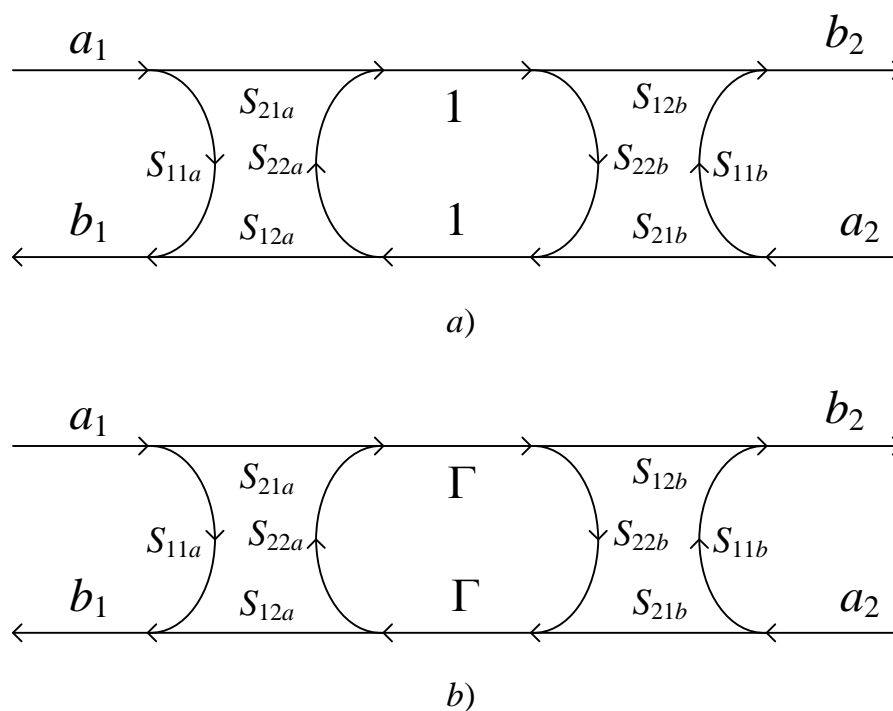


Рис. 2. Каскадное соединение цепей А и В:

a) в направляющих системах,

b) в открытом пространстве

В случае L-соединения (рисунок 2а) в направляющих системах цепи А и В соединены между собой напрямую, что можно интерпретировать как трехкаскадное соединение с единичной матрицей рассеяния. В открытом пространстве трехкаскадная модель L-соединения (рисунок 2б) имеет вместо единичной матрицы матрицу с коэффициентами передачи, равными коэффициенту отражения Γ от плоского отражателя. Параметры рассеяния цепей А и В следует определять из соотношений [8]:

$$\begin{aligned}
S_{22}^a &= \Gamma \left(\frac{S_{12}^a}{S_{12}^b} \right) \frac{S_{22}^L - S_{11}^b}{S_{12}^L}, & S_{22}^b &= \Gamma \left(\frac{S_{12}^b}{S_{12}^a} \right) \frac{S_{11}^L - S_{11}^a}{S_{21}^L}, \\
S_{21}^a S_{12}^a &= \frac{S_{21}^L}{\Gamma} \left(\frac{S_{12}^a}{S_{12}^b} \right) (1 - S_{22}^a S_{22}^b), & S_{21}^b S_{12}^b &= \frac{S_{12}^L}{\Gamma} \left(\frac{S_{12}^b}{S_{12}^a} \right) (1 - S_{22}^a S_{22}^b), \\
S_{21}^a S_{12}^b &= \frac{S_{21}^L}{\Gamma} (1 - S_{22}^a S_{22}^b), & S_{21}^b S_{12}^a &= \frac{S_{12}^L}{\Gamma} (1 - S_{22}^a S_{22}^b).
\end{aligned} \tag{1}$$

Измерения в открытом пространстве проводят для двух поляризаций зондирующих электромагнитных волн: - вертикальной и горизонтальной. Вертикальная поляризация вектора напряженности электрического поля дает так называемую перпендикулярную поляризацию [9], у которой $\Gamma_{\perp} = -1$, горизонтальная поляризация – параллельную поляризацию с $\Gamma_{\parallel} = 1$.

В отличие от LR-метода в направляющих системах, модули коэффициентов передачи S_{21}^a , S_{21}^b цепей А и В в прямом направлении в открытом пространстве близки к единице, а модули обратных коэффициентов передачи S_{12}^a , S_{12}^b имеют весьма малое значение. Этот эффект, как показано на рисунке 1, объясняется формированием расходящегося луча передающей антенны, освещающего зеркало. Если зеркало имеет достаточно большой размер, то практически вся энергия электромагнитного излучения, сосредоточенная в главном лепестке диаграммы направленности передающей антенны, отразится в направлении приемной антенны и модуль коэффициента передачи S_{21}^a будет близок к единице. Отраженное от зеркала электромагнитное излучение в области приемной антенны будет освещать площадку существенно большего размера по отношению к площади зеркала. Модуль обратного коэффициента передачи S_{12}^b можно приближенно оценить отношением площади апертуры приемной антенны к площади виртуальной площадки, освещенной в плоскости приемной антенны. Это отношение будет много меньше единицы.

Таким образом, цепи А и В в случае открытого пространства будут невзаимными. Тем не менее, следует отметить тот факт, что невзаимные свойства цепей ни как не отразятся на работоспособности предлагаемого метода измерения в открытом пространстве, поскольку нет необходимости выделять прямые S_{21}^a, S_{21}^b и обратные S_{12}^a, S_{12}^b коэффициенты передачи из их произведений: - $S_{21}^a S_{12}^a$, $S_{21}^b S_{12}^b$, $S_{21}^a S_{12}^b$.

В отличие от LR и экспресс методов [8] коэффициенты отражения от антенн со стороны их коаксиальных разъемов определяют не как средние линии квазипериодических комплексных функций частоты электромагнитного поля. Коэффициенты отражения S_{11}^a и S_{11}^b получают прямыми измерениями коэффициентов отражения от М-соединения, в котором отражающую пластину поворачивают на угол порядка 45° , как показано на рисунке 3, чем обеспечивается режим идеального согласования антенны с открытым пространством.

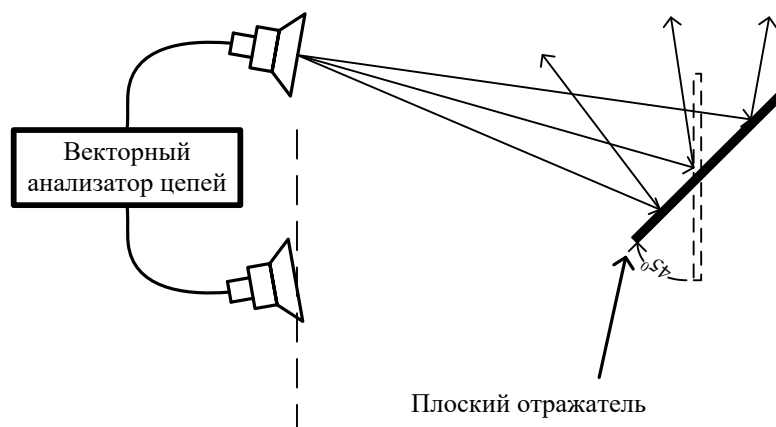


Рис. 3. М-соединение антенн в открытом пространстве

Ввиду малого, порядка 50 дБ, динамического диапазона измерений коэффициентов отражения анализаторами цепей, по сравнению с широким, порядка 140 дБ, диапазоном измерений коэффициентов передачи, информацию об электромагнитных характеристиках объектов в открытом пространстве, следует получать только из результатов измерений коэффициентов передачи. Поэтому для измерения электромагнитных характеристик материалов, поглощающих покрытий и ЭПР малоразмерных объектов предлагается метод дистанционного контроля бистатического коэффициента отражения в открытом пространстве. Причем бистатический коэффициент отражения Γ_{dut} , содержащий информацию о характеристике плоского образца или малоразмерного объекта при наклонном падении зондирующей электромагнитной волны, вычисляют по коэффициенту передачи S_{21} в широком динамическом диапазоне из соотношения:

$$\Gamma_{dut} = \frac{S_{21} S_{12}^a S_{21}^b}{S_{22}^a (S_{22}^b \det S - S_{11} \det S^b) + (\det S^b - S_{22} S_{22}^b) \det S^a} \quad (2)$$

Существенно, что коэффициенты отражения от плоских образцов неметаллических материалов можно измерять с отражающей пластиной и без пластины при вертикальной и горизонтальной поляризации электромагнитного поля. Такая особенность измерений объектов в открытом пространстве позволяет оценивать достоверность получаемых результатов. Что же касается отношения коэффициентов передачи S_{12}^a/S_{12}^b , то, как выяснилось в результате исследования, эта величина не влияет на результаты измерений в открытом пространстве и может быть принята равной 1.

Принципиальная особенность предлагаемого ЛМ-метода измерений состоит в исключении влияния параметров антенн и окружающего пространства на получаемые результаты. Измерения следует проводить при вертикальной и горизонтальной поляризации зондирующей электромагнитной волны в широком диапазоне частот, чтобы обнаружить возможные резонансные пики значений модулей коэффициентов отражения от контролируемых объектов на частотах, при которых элементы их конструкции будут кратны длине λ зондирующей электромагнитной волны. Для изменения поляризации с вертикальной на горизонтальную антенны поворачивают на угол 90° , не изменяя ориентации в пространстве их оптических осей. Влияние не прямых трасс при зондировании малоразмерных объектов направленными антеннами в обычном лабораторном помещении маловероятно. Оно все же может проявиться в возникновении квазигармонических колебаний или шумовой дорожки около воображаемых средних линий частотно зависимых результатов обработки данных, несущих информацию о характеристиках контролируемых объектов.

Для реализации LM-метода контроля коэффициентов отражения от плоских образцов материалов и поглощающих покрытий, а также бистатической ЭПР малоразмерных объектов разработан и испытан компактный антенный полигон с наклонным падением зондирующей электромагнитной волны. Для экспериментальной отладки методик измерения были использованы сравнительно узкополосные антенны спутникового телевидения (диапазон частот 10,7 — 12,75 ГГц). Такая антенна представляет собой систему из параболического зеркала диаметром 55 см и рупорной антенны на основе круглого волновода со встроенным преобразователем частоты. При доработке антенны преобразователь частот был удален с заменой на штыревой волноводно-коаксиальный переход. Для связи с анализатором цепей установлен SMA разъем.

Кроме того, проводились эксперименты с многоэлементной фазированной антенной решеткой (ФАР). Она построена на плоских печатных волноводах. Восемь таких плат объединяются в систему с 64 излучателями. Распределение энергии между платами производится делителем мощности на плоских печатных волноводах. Частотный диапазон данной системы составляет 8 – 11 ГГц.

Измерения проводились с двумя видами поляризаций: вертикальной и горизонтальной. На рисунке 4 антенные системы настроены на вертикальную поляризацию. Для изменения поляризации в зеркально-рупорных антеннах производился поворот рупоров, а в системе с ФАР – самих антенн.

Для калибровки системы использовались пространственные аналоги отрезка линии передачи (калибровка на проход) и согласованной нагрузки. Обе меры реализованы путем установки на пути распространения электромагнитной волны отражателя в виде металлического листа, размеры которого превышают размеры главного лепестка диаграммы направленности используемых антенн.

При калибровке на проход отражатель устанавливается строго вертикально, обеспечивая попадание отраженной волны в приемную антенну. Для реализации согласования отражатель отклонялся от вертикального положения на определенный угол. При этом отраженная волна не попадает в приемную антенну. Можно подобрать такой угол наклона, при котором и вторичные отражения от элементов конструкции помещения не будут влиять на результат измерения. Этот угол можно подобрать экспериментально по результатам измерения коэффициента передачи (максимальная величина затухания и отсутствие высокочастотных колебаний АЧХ).



Рис. 4. Антенная система с ФАР

При определении диэлектрических параметров листовых материалов (полиметилметакрилат) производились два измерения: с отражателем, установленным за материалом и с согласованием прошедшей через материал волны (с наклонным расположением отражателя).

Объектами экспериментальных исследований были пластины различного диаметра из алюминиевого сплава D16: - $250 \times 10 \times 2$ мм, $250 \times 20 \times 2$ мм, $250 \times 40 \times 2$ мм.

Распределение амплитуды электромагнитного поля зондирующей электромагнитной волны в плоскости позиционирования контролируемых объектов было восстановлено по результатам измерения пластин из сплава D16, подвешенных вертикально к устройству позиционирования.

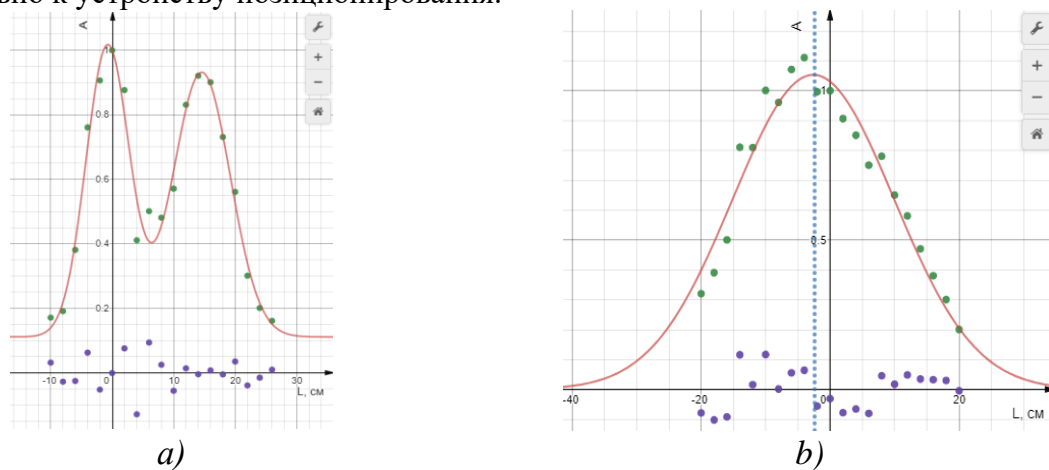


Рис. 5. Распределение амплитуды падающей волны по горизонтали:
а) – зеркальных антенн, б) – фазированных антенных решеток (измерения – зеленый цвет, аппроксимация Гауссовой функцией – красный, невязка – синий).

Результаты измерения распределения по горизонтали амплитуды электромагнитного поля зондирующих электромагнитных волн в плоскости позиционирования контролируемых объектов, полученные с помощью зеркальных антенн и фазированных антенных решеток, приведены на рисунке 5. Расстояние от зеркальных антенн до плоскости позиционирования объектов 2.07 метра, от ФАР до объектов – 2.60 метра. Распределения амплитуды электромагнитного поля зондирующей электромагнитной волны в плоскости позиционирования контролируемых объектов хорошо описываются Гауссовой функцией.

В силу острой направленности зеркальных антенн, в распределении наблюдаются ярко выраженные два максимума. Очевидно, что, более точной ориентацией оптических осей антенн с помощью промышленных средств позиционирования, можно получить плоский участок на вершине распределения.

Таким образом, эффективность предлагаемого решения экспериментально подтверждена на примере измерения коэффициентов отражения от металлических отражателей разной площади.

Библиографический список

1. Андреев И.Л. Частотно-временной метод коррекции результатов измерения объемных диаграмм направленности антенн / И.Л. Андреев, А.В. Кузьма // Измерительная техника. 1996. – № 12. – С. 44-47.
2. Levitas B. N. Time Domain Antenna Measurement Systems / B.N. Levitas// Ultrawideband and Ultrashort Impulse Signals, 18-22 September, 2006, Sevastopol, Ukraine, pp. 90-95.
3. Bartik H. Antenna Measurements Using the Mirror Method with Gating in a Time Domain / N. Bartik. // Radioengineering, Vol. 14, No. 4, December 2005, pp.58-62.

4. R. K de Jongh. Antenna Time-Domain Measurement Techniques /R. K de Jongh, M. Hajian, L. P. Ligthart // IEEE Antennas and Propagation Magazine, Vol. 39, No. 5, October 1997, pp. 7-12.
5. Lima D.R. Accuracy of GRL Calibration Considering Time Domain Gating for the Calculation of Permittivity Parameter in Free Space Technique / Lima D.R., Junqueira C., Canto M.A.R., Migliano A.C.// Proc. of Microwave and Optoelectronics Conference (IMOC), Brazil, 3-6 Nov, 2015, pp. 1-5.
6. Khaleghi A. Time-Domain Measurement of Antenna Efficiency in Reverberation Chamber / Khaleghi A. // IEEE Transactions on Antennas and Propagation, Vol. 57, No. 3, March 2009, pp. 817-821.
7. Костенко А.А. Квазиоптика: исторические предпосылки и современные тенденции развития /А.А. Костенко // Радиофизика и радиоастрономия, 2000, т. 5, №3, стр. 221-246
8. Евсеев В.И. LR и экспресс-метод контроля параметров электронных компонентов в полосковых линиях передачи / В.И. Евсеев, С.М. Никулин // Электроника и микроэлектроника СВЧ. 2020, с. 18 -22 ISBN978-5-7629-2456-6.
9. Никольский В.В. Электродинамика и распространение радиоволн. В.В. Никольский. Учебное пособие. Главная редакция физико-математической литературы изд-ва «Наука», М., 1978, 544 с.

V.I. Evseev, E.A. Lupanova

Nizhny Novgorod state technical university n.a. R.E. Alekseev

The paper considers the 2L Express method for measuring the S-parameters of devices, which is a development of the LRT, LR and Express methods. The proposed method solves the problem of determining the midline, which is the "thin" place of these methods. The LM-method for monitoring the electromagnetic characteristics of materials, absorbing coatings and the bistatic effective scattering area of small objects is discussed. The fundamental feature of the proposed solution in comparison with the existing methods of antenna measurements is to exclude the influence on the results obtained from the characteristics of the transmitting and receiving antennas, including the surrounding spaces between the antennas and the controlled object. In addition, the method makes it possible to exclude the influence of indirect paths of propagation of electromagnetic waves.

Keywords: free space, electromagnetic characteristics of materials, bistatic effective scattering area of small objects

РАЗРАБОТКА И ОПТИМИЗАЦИЯ МОДУЛЯ ЦИФРОВОГО ДИАГРАММООБРАЗОВАНИЯ ДЛЯ АВТОМОБИЛЬНОГО РАДАРА

Нижегородский государственный технический университет им. Р.Е. Алексеева

Рассмотрены вопросы разработки модуля цифрового диаграммообразования на базе ПЛИС для применения в автомобильном радаре, и проведен анализ путей оптимизации его архитектуры с целью снижения затрачиваемых ресурсов и улучшения временных характеристик.

Ключевые слова: автомобильный радар, цифровое диаграммообразование, формирование лучей, ПЛИС, оптимизация ресурсов

Введение

Одной из основных задач, решаемых автомобильным радаром в составе системы интеллектуальной помощи водителю (ADAS – advanced driver-assistance systems) является определение азимутального положения объектов относительно автомобиля и разрешение обнаруженных объектов по азимуту. Для решения этой задачи в радарх применяются фазированные антенные решетки миллиметрового диапазона в совокупности с подсистемой цифрового диаграммообразования. Данная подсистема позволяет сформировать набор диаграмм направленности антенны с ярко выраженными основными лепестками в направлениях, отличающихся на единицы градусов. Каждую такую диаграмму направленности в зарубежной литературе принято называть лучом (beam), а процесс формирования набора диаграмм – формированием лучей (beam forming) [1].

Формируемый набор диаграмм направленностей (лучей) должен удовлетворять техническим требованиям, предъявляемым к радару, в части обеспечения перекрытия необходимого диапазона азимутальных углов (угла обзора) и заданной точности определения азимутального положения обнаруживаемых объектов и разрешающей способности по азимуту.

Ведущие мировые производители автомобильных радаров в качестве аппаратной платформы, на базе которой производится цифровая обработка сигналов, включая формирование лучей, используют либо специализированные микропроцессоры с аппаратными ускорителями, либо программируемые логические интегральные схемы (ПЛИС) типа FPGA (Field Programmable Gate Array) с встроенными специализированными блоками (оперативная память, умножители).

В данной статье рассматриваются вопросы разработки модуля цифрового диаграммообразования (формирования лучей) на базе ПЛИС для применения в автомобильном радаре и проводится анализ путей оптимизации его архитектуры с целью снижения затрачиваемых ресурсов и улучшения временных характеристик (увеличения максимально возможной тактовой частоты работы схемы).

Общая постановка задачи

Автомобильный радар, являющийся целевым устройством для разрабатываемого формирователя лучей, имеет антенную решетку, состоящую из $K = 12$ элементов. Сигнал от каждого элемента обрабатывается отдельно сначала в аналоговом, а затем в цифровом виде.

Все пространство обзора рассматриваемого радара разбивается на две части: ближняя и дальняя зоны. Отличаются эти зоны диапазоном углов обзора. В ближней зоне необходимо слежение не только за транспортными средствами на проезжей части перед автомобилем, но и за объектами на обочине (пешеходы, велосипедисты), что требует обеспечение широкого угла обзора и относительно большого числа лучей $Q_{БЗ} = 11$. В дальней зоне приоритет отдается транспортным средствам и препятствиям, находящимся

на пути следования автомобиля. Здесь требуется обеспечение более узкого угла обзора, чем в ближней зоне, и меньшего числа лучей $Q_{дз} = 9$.

Сканирование ближней и дальней зоны в рассматриваемом радаре производится последовательно. При этом вводится понятие режима работы радара – режим ближнего действия и режим дальнего действия. Отличаются эти режимы уровнем излучаемой мощности и процессом формирования лучей. Таким образом, подсистема диаграммообразования автомобильного радара должна иметь в своем составе либо одно двухрежимное исполнительное устройство, либо два отдельных исполнительных устройства для каждого из режимов.

Цифровая обработка рассматриваемого радара реализуется на базе ПЛИС10CL120YF48417 типа FPGАсемейства Cyclon10LP фирмы Intel [2, 3]. Вычисления производятся в формате с фиксированной точкой. Параметры обработки:

- частота дискретизации – $f_s = 12$ МГц;
- вид входных данных – комплексные отсчеты в формате с фиксированной точкой;
- разрядность входных данных – 25;
- разрядность весовых коэффициентов – 17;
- разрядность выходных данных – 27.

Таким образом, разрабатываемый формирователь лучей должен представлять собой программный модуль, написанный на языке описания аппаратуры (hard ware description language – HDL) и поддерживающий возможность его имплементации в ПЛИС, указанного семейства.

При реализации устройств на базе ПЛИС одними из ключевых являются задачи минимизации объема затрачиваемых аппаратных ресурсов, обеспечивающая допустимое ухудшение качества обработки (либо не допускающая его вовсе), и обеспечения требуемого быстродействия. Как правило, результатом их решения является компромиссный вариант, обеспечивающий баланс между объемом аппаратных затрат и минимально необходимой тактовой частоты работы схемы.

При разработке зачастую приходится отказываться от простых и очевидных решений, в которых имплементации подлежит каждый элементарный функциональный блок, в пользу более сложных вариантов, когда по возможности одни и те же элементарные блоки используются многократно в одном цикле работы реализуемого устройства (например, при вычислении одного выходного отсчета). Такие варианты обеспечивают экономию аппаратных ресурсов, но требуют повышения тактовой частоты работы схемы, что может оказаться проблематичным ввиду наличия некоторой верхней границы быстродействия, определяемой особенностями архитектуры выбранного семейства ПЛИС. Особенно это касается используемого семейства Cyclon10LP, позиционируемого как семейство ПЛИС с низким энергопотреблением (lowpower – LP) и характеризующегося в связи с этим невысокой производительностью.

Анализ возможных вариантов реализации

Формирование лучей (ФЛ) математически можно представить, как весовое суммирование отсчетов K каналов антенной решетки:

$$Y_q(n) = \sum_{k=1}^K X_k(n) \cdot W_{q,k}, \quad (1)$$

где n – номер отсчета; k – номер канала антенной решетки ($k = 1 \dots K$); K – число обрабатываемых каналов антенной решетки; q – номер луча ($q = 1 \dots Q$); Q – общее число лучей; $Y_q(n)$ – n -й выходной комплексный отсчет q -го луча; $X_k(n)$ – n -й входной комплексный отсчет k -го канала антенной решетки, $W_{q,k}$ – комплексный весовой коэффициент для k -го канала антенной решетки, используемый при получении отсчета q -го луча.

На рис. 1 представлена функциональная структура алгоритма формирования одного (q -го) луча.

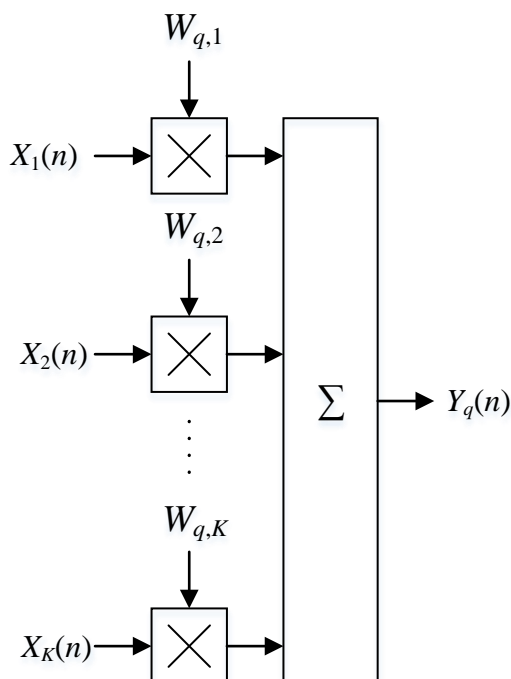


Рис. 1. Функциональная структура алгоритма формирования одного q -го луча

Самым простым и очевидным вариантом реализации формирователя лучей является выполнение его в виде нескольких параллельно включенных наборов функциональных блоков, изображенных на рис.1 (K комплексных умножителей и один параллельный сумматор). При таком подходе данные всех каналов антенной решетки подаются на вход формирователя одновременно (параллельно), выходные отсчеты формируются также одновременно. Результирующая задержка появления выходных данных относительно момента поступления входных отсчетов будет равна сумме конвейерных задержек одного умножителя и одного многовходового сумматора.

Преимуществом данного варианта является возможность синхронной работы устройства с тактовой частотой f_{clk} , равной частоте поступления входных отсчетов, то есть с частотой дискретизации f_s , равной, как было указано выше, 12 МГц.

Для обеспечения возможности работы формирователя лучей в режимах дальнего и ближнего действия целесообразно использовать одни и те же вычислительные ресурсы (умножители и сумматоры) при использовании отдельных наборов весовых коэффициентов $W_{q, k}$ для каждого из режимов. Разные наборы можно использовать, например, путем хранения коэффициентов в разных банках встроенных блоков оперативной памяти и переключения банков в зависимости от режима. При этом все вычислительные ресурсы и объем необходимой памяти должны планироваться исходя из максимальных требований каждого режима. Так определяющим параметром в вопросе определения затрачиваемых вычислительных ресурсов является число лучей Q . В зависимости от режима работы — это число может составлять, как было указано выше, либо $Q_{БЗ} = 11$, либо $Q_{ДЗ} = 9$. Очевидно, что решающую роль в планировании ресурсов будет играть максимальное число лучей:

$$Q_{max} = \text{MAX}\{Q_{БЗ}, Q_{ДЗ}\} = \text{MAX}\{11,9\} = 11. \quad (2)$$

При таком способе реализации формирователь лучей будет содержать $K \times Q_{max} = 12 \times 11 = 132$ комплексных умножителей и $Q_{max} = 11$ комплексных многовходовых

сумматора или, что то же самое, $Q_{max} \times (K - 1) = 11 \times (12 - 1) = 121$ комплексных двухвходовых сумматора. Принимая во внимание, что каждое комплексное умножение выполняется с помощью четырех операций умножения и двух операций сложения действительных чисел:

$$(a + jb) \times (c + jd) = (a \times c - b \times d) + j(a \times d + b \times c),$$

а каждое комплексное сложение выполняется с помощью двух операций сложения действительных чисел, можно получить оценку необходимых вычислительных ресурсов:

- $K \times Q_{max} \times 4 = 12 \times 11 \times 4 = 528$ умножителей;
- $Q_{max} \times (K - 1) \times 2 + K \times Q_{max} \times 2 = 11 \times (12 - 1) \times 2 + 12 \times 11 \times 2 = 506$ сумматоров.

Для реализации умножения в ПЛИС целесообразно использовать встроенные специализированные блоки умножителей (DSP-блоки), предназначенные для работы с операндами разрядностью 9 или 18. Не вдаваясь в детали специфики использования аппаратных умножителей ПЛИС с операндами разрядностью выше 18, можно сказать, что рассматриваемый вариант является не реализуемым в случае использования указанной выше микросхемы 10CL120YF484I7, поскольку согласно [2, 3] эта микросхема имеет всего лишь 288 DSP-блоков, обеспечивающих работу с операндами разрядностью 18, в то время как требуется как минимум 528 таких блоков.

Таким образом, применение параллельной схемы построения обработки является невозможным в существующих условиях. Требуется применение последовательно-параллельного способа организации вычислений. Особенности такого способа является снижение аппаратных затрат за счет многократного использования одних и тех же вычислительных ресурсов в течение одного цикла работы устройства и повышения требований к его производительности (минимальной тактовой частоте). Можно рассмотреть два варианта реализации такого способа – при параллельном поступлении данных на вход и при последовательном поступлении.

Обобщенная структура формирователя лучей для последовательно-параллельного способа организации вычислений при параллельном поступлении данных на вход приведена на рис. 2. Здесь IData1...IData12 – входные шины данных, содержащие комплексные отсчеты сигналов каналов антенной решетки; OData1...OData11 – выходные шины данных, содержащие выходные комплексные отсчеты лучей; ROM1...ROM12 – блоки памяти, хранящие значения коэффициентов $W_{q, k}$; RG – регистры для хранения выходных комплексных отсчетов.

Логика работы данной схемы заключается в следующем. На вход формирователя одновременно поступают значения отсчетов сигналов $K = 12$ каналов антенной решетки, после чего они не меняются в течение 11 периодов тактового сигнала (тактов). За это время производится вычисление 11 значений весовых сумм согласно (1), каждое из которых является отсчетом одного луча. При этом необходимые значения весовых коэффициентов $W_{q, k}$ вычитываются из блоков памяти ROM. Результат вычисления каждой суммы записывается в свой выходной регистр RG. За своевременную выдачу на вход умножителей соответствующих коэффициентов из блоков памяти ROM и своевременное разрешение записи в конкретный выходной регистр RG отвечает устройство управления. Возможность использования разных наборов коэффициентов в разных зонах (ближней/дальней) реализуется с помощью увеличенного объема памяти блоков ROM и выдачей корректных сигналов управления на них устройством управления в зависимости от обрабатываемой зоны (используются разные диапазоны адресов блоков памяти при обработке разных зон).

Особенностью данного варианта является необходимость работы схемы на повышенной тактовой частоте, которая должна быть выше частоты дискретизации не менее чем в Q_{max} раз, то есть должна составлять как минимум $f_{clk} = f_s \times Q_{max} = 12 \times 11 = 132$ МГц.

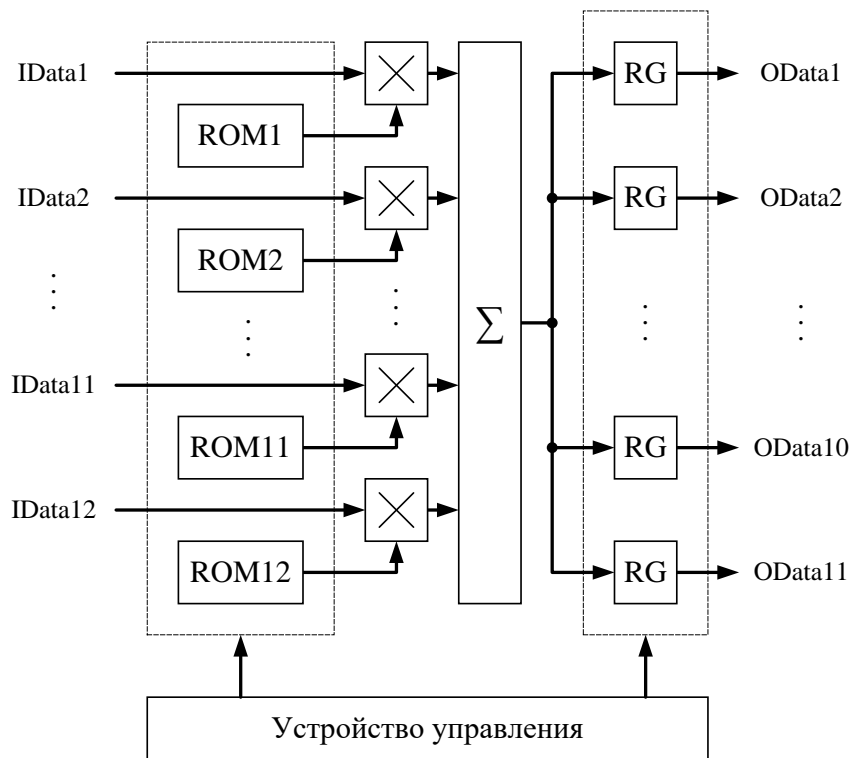


Рис. 2. Обобщенная структура формирователя лучей при параллельном поступлении данных на вход

Следуя рассуждениям, использованным при рассмотрении параллельной схемы построения обработки, можно получить общую оценку необходимых вычислительных ресурсов для рассматриваемого варианта:

- $K \times 4 = 12 \times 4 = 48$ умножителей;
- $(K - 1) \times 2 + K \times 2 = (12 - 1) \times 2 + 12 \times 2 = 46$ двух входовых сумматора.

Как видно объем ресурсов для данного варианта значительно ниже, чем для варианта с полностью параллельной структурой. Платой за снижение аппаратных затрат является повышение требований к быстродействию схемы. Однако на основании опыта работы с семейством Cyclon10LP можно утверждать, что требование минимальной тактовой частоты $f_{clk} = 132$ МГц является вполне приемлемым и достижимым для схем, реализованных на базе ПЛИС данного семейства.

Обобщенная структура формирователя лучей для последовательно-параллельного способа организации вычислений при последовательном поступлении данных на вход приведена на рис. 3. Здесь *IData* – входная шина данных; *OData* – выходная шина данных; ROM1...ROM11 – блоки памяти, хранящие значения коэффициентов $W_{q, k}$; RG – регистры для временного хранения выходных комплексных отсчетов (на время выдачи их содержимого на выходную шину *OData*).

Данная схема функционирует следующим образом. На вход формирователя по шине *IData* в течение 12 тактов последовательно поступают значения отсчетов сигналов $K = 12$ каналов антенной решетки. Входные данные поступают на комплексные умножители совместно с комплексными коэффициентами $W_{q, k}$, которые вычитываются из блоков памяти ROM. В течение 12 тактов с помощью умножителей и аккумуляторов параллельно производится вычисление $Q_{max} = 11$ весовых сумм. Их результирующие значения записываются в регистры RG, после чего выдаются на выходную шину *OData* с помощью мультиплексора MUX.

За своевременную подачу на вход умножителей соответствующих коэффициентов из блоков памяти ROM, сброс аккумуляторов, разрешение записи в регистры RG и

переключение выходного мультиплексора MUX отвечает устройство управления. Как и в предыдущем варианте, возможность использования разных наборов коэффициентов в разных зонах (ближней/дальней) реализуется с помощью специальной организации адресного пространства памяти блоков ROM.

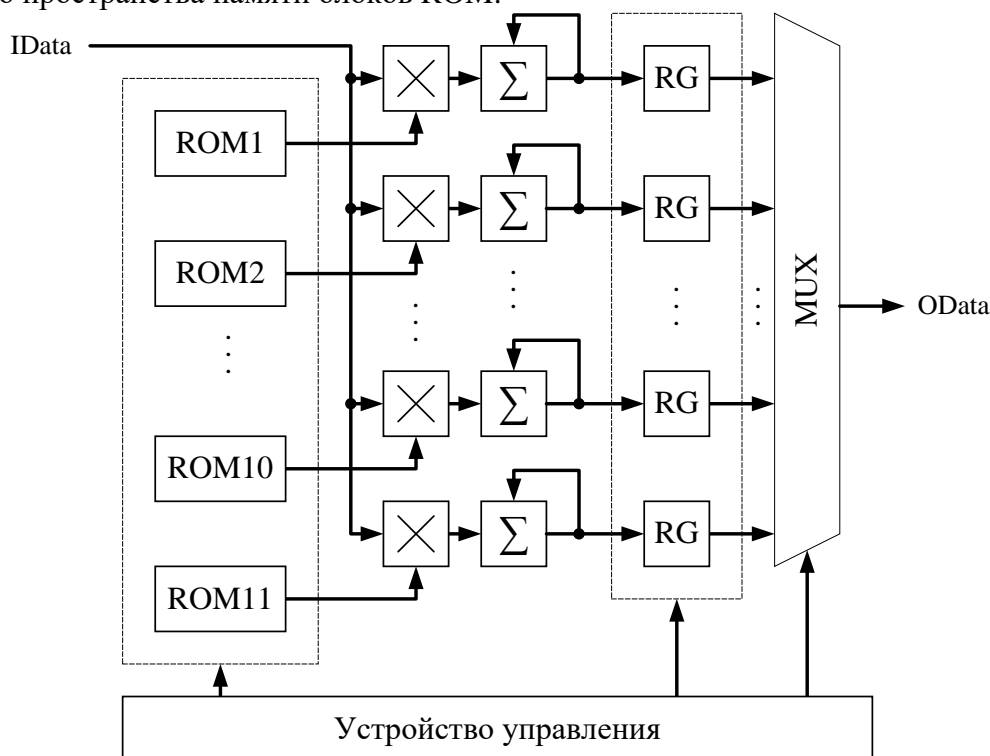


Рис. 3. Обобщенная структура формирователя лучей при последовательном поступлении данных на вход

У данного варианта реализации также есть требование работы на повышенной тактовой частоте, которая должна быть выше частоты дискретизации не менее чем в K раз $-f_{clk} = f_s \times K = 12 \times 12 = 144$ МГц.

Следуя рассуждениям, использованным ранее, можно получить общую оценку необходимых вычислительных ресурсов для рассматриваемого варианта:

- $Q_{max} \times 4 = 11 \times 4 = 44$ умножителей;
- $Q_{max} \times 2 = 11 \times 2 = 22$ двухвходовых сумматора;
- $Q_{max} \times 2 = 11 \times 2 = 22$ аккумулятора (двухвходовых сумматора).

Как видно объем вычислительных ресурсов, необходимый для реализации данного варианта, несколько ниже, чем для варианта с параллельным поступлением входных данных (рис. 2). Однако требование по минимальной тактовой частоте здесь выше (144 МГц против 132 МГц).

Исходя из полученных результатов, можно сделать вывод, что наиболее предпочтительным по критерию минимизации вычислительных затрат является второй вариант (рис. 3).

Следует отметить, что в случае такой конфигурации системы, когда число лучей Q_{max} превышает число входных каналов K , наименее затратным по ресурсам и наиболее требовательным по частоте станет первый вариант с параллельным поступлением данных на вход (рис. 2).

Оптимизация комплексных умножителей

Особенностью разрабатываемого формирователя лучей является достаточно высокая разрядность входных данных – 25 при типовой разрядности операндов DSP-блоков ПЛИС, равной 18. В этом случае одно действительное умножение не может быть выполнено на одном DSP-блоке – требуется специальная схема каскадирования нескольких таких блоков. В случае использования этой схемы операция умножения двух операндов A и B представляется следующим образом:

$$C = A \times B = (A_H \times 2^{WL} + A_L) \times B = A_H \times B \times 2^{WL} + A_L \times B, \quad (3)$$

где A_H – старшая часть операнда A (занимающая старшие $WA - WL$ разрядов); A_L – младшая часть операнда A (занимающая младшие WL разрядов); WA – разрядность операнда A ; WL – разрядность младшей части операнда A . Здесь предполагается, что разрядность операнда A не превышает $18 \times 2 = 36$, а разрядность операнда B не превышает 18.

Таким образом, одно действительное умножение реализуется в виде суммы двух произведений: $A_H \times B$ и $A_L \times B$. Операция умножения результата произведения $A_H \times B$ на значение 2^{WL} не требует вычислительных затрат, поскольку сводится к простому дополнению нулями этого результата со стороны младших разрядов.

Процесс выполнения операции умножения согласно (3) иллюстрирует рис. 4. Здесь WB – разрядность операнда B ; Z – знаковый разряд.

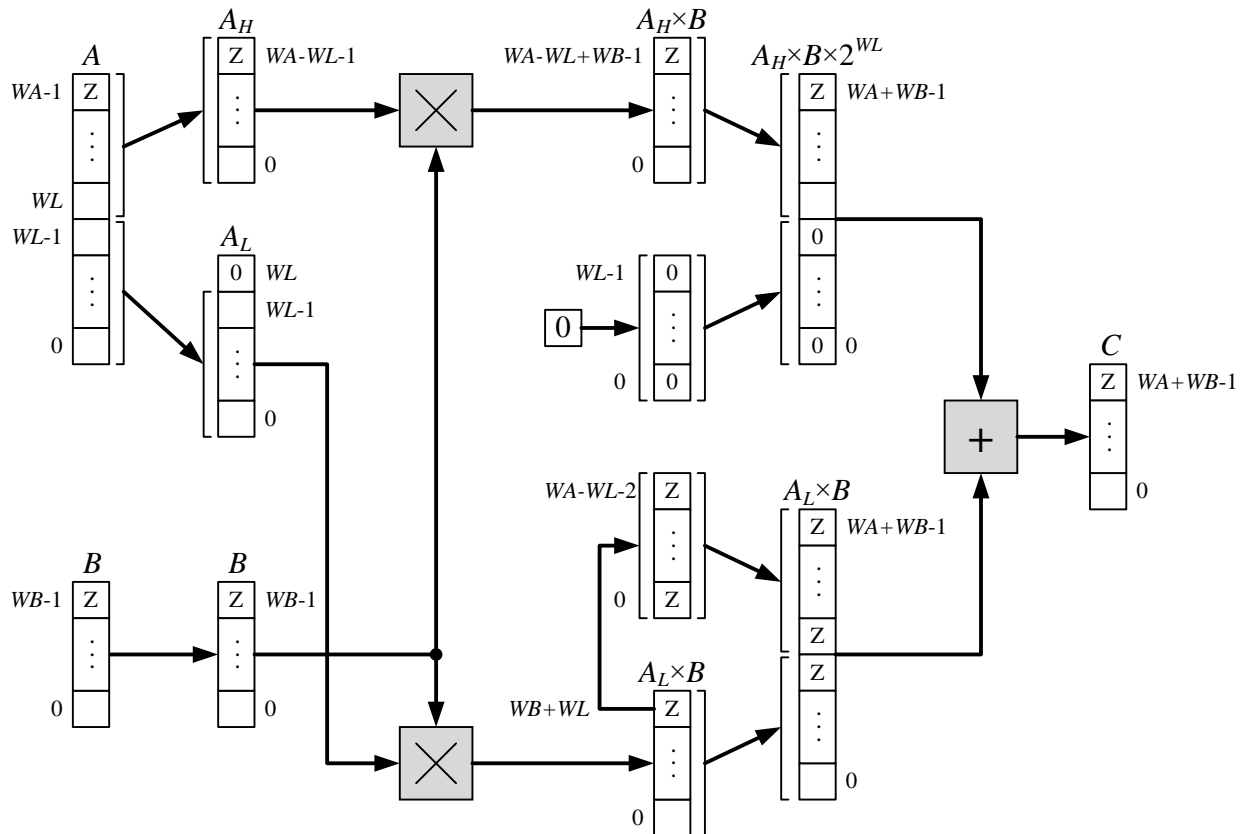


Рис. 4. Процесс выполнения операции умножения при использовании схемы каскадирования умножителей

Необходимость применения описанной схемы приводит к двукратному увеличению числа DSP-блоков, требуемых для реализации формирователя лучей. Кроме того, ввиду особенностей размещения DSP-блоков на кристалле ПЛИС, снижается быстродействие всей схемы за счет увеличения длин комбинационных путей в окрестности умножителей.

При решении проблемы снижения быстродействия необходимо обращать внимание на сигнальные пути на входе и выходе умножителей. Именно эти пути являются наиболее критичными в плане вносимой задержки. Причин здесь две. Во-первых, операция умножения является сама по себе затратной по времени. Во-вторых, расположение DSP-блоков на кристалле является фиксированным, в то время как прочая логика (например, сумматор, изображенный на рис. 4), реализуемая на базе конфигурационных логических блоков (распределенная логика), может располагаться на кристалле в произвольных местах в зависимости от результатов работы трассировщика САПР Quartus. В итоге без принятия мер по включению в тракты умножителей синхронных триггеров длина комбинационных путей может оказаться чрезмерно большой, что не позволит обеспечить выполнение требования по быстродействию.

Наилучшим решением в данной ситуации является использование синхронных триггеров из состава DSP-блоков, как на выходе, так и на входе умножителей. В этом случае длинный комбинационный путь, проходящий через DSP-блок, будет разбит на три части:

- от распределенной логики до DSP-блока (входной сигнальный путь);
- внутри DSP-блока (внутренний сигнальный путь);
- от DSP-блока до распределенной логики (выходной сигнальный путь).

Однако при реализации устройства с использованием языков описания аппаратуры (VHDL, Verilog) можно столкнуться проблемой – САПР Quartus автоматически не задействует входной триггер DSP-блоков. На рис. 5 приведена структура DSP-блока с указанием задействованных элементов, характеризующая описанную ситуацию. Данная структура получена с помощью инструмента Resource Property Editor из состава САПР Quartus. Здесь неиспользуемые элементы отмечены светло-серым цветом.

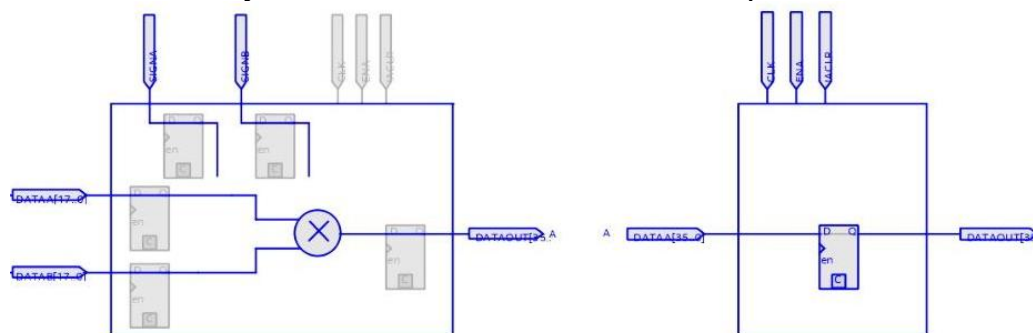


Рис. 5. Структура DSP-блока без использования входных триггеров

Выходом из сложившейся ситуации является использование умножителей из набора блоков интеллектуальной собственности (IP-ядер), поставляемого в составе САПР Quartus. В настройках IP-ядра умножителя можно задать значение задержки (latency) больше двух, что приведет к использованию входных триггеров (рис. 6).

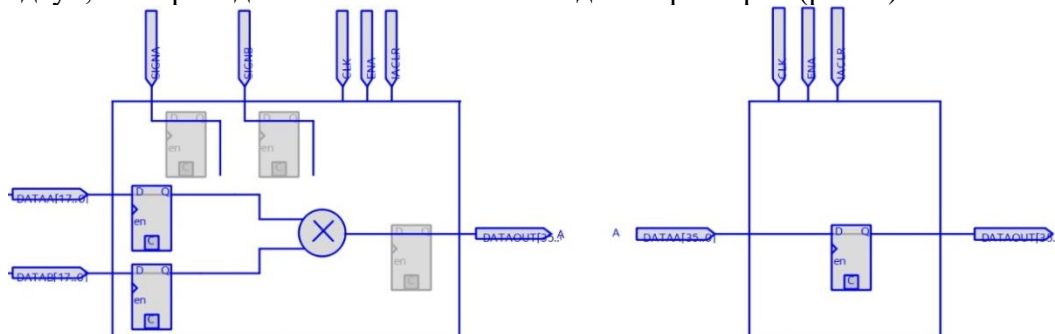


Рис. 6. Структура DSP-блока при использовании входных триггеров

Для снижения числа DSP-блоков, используемых для реализации формирователя лучей, можно применить подход, описанный в [4], позволяющий уменьшить число операций умножения действительных чисел, приходящихся на одно комплексное умножение:

$$(a + jb) \times (c + jd) = (a - b) \times d + (c - d) \times a + j(a - b) \times d + j(c + d) \times b = e + jf.$$

Как видно из приведенной формулы при вычислении реальной и мнимой части результата произведения используется одно и то же произведение $(a - b) \times d$, которое можно вычислить всего один раз. Таким образом, число операций умножения можно снизить с четырех до трех. Однако при этом будет увеличено число операций сложения и вычитания. Функциональная структура оптимизированного комплексного умножителя представлена на рис. 7.

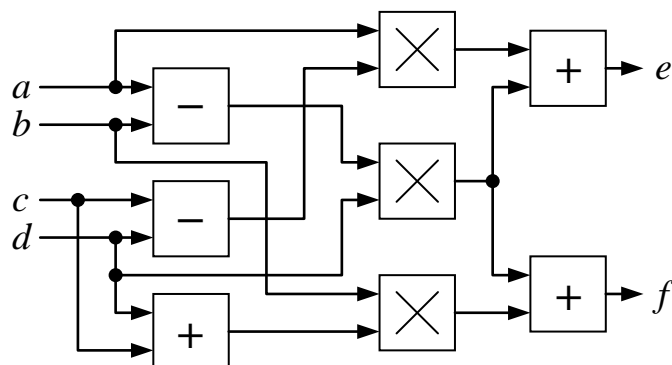


Рис. 7. Функциональная структура алгоритма комплексного умножения

В приведенной схеме присутствуют три дополнительные операции сложения/вычитания. Однако их можно уменьшить, если операции $(c - d)$ и $(c + d)$ вычислять заранее и использовать их результат на правах значений весовых коэффициентов $W_{g, k}$. Это приведет к некоторому увеличению объема используемой для хранения коэффициентов встроенной памяти на ПЛИС, но ввиду относительно небольшого их числа, такое увеличение не будет оказывать большого влияния.

С учетом всего вышесказанного можно получить итоговую оценку числа DSP-блоков, требуемых для реализации рассмотренных ранее вариантов построения формирователя лучей. Так для реализации варианта с параллельным поступлением данных на вход (рис. 2) потребуется: $K \times 3 \times 2 = 12 \times 3 \times 2 = 72$ DSP-блока. Здесь учитывается использование трех операций действительного умножения на одну операцию комплексного умножения и необходимость использования двух DSP-блоков для выполнения операции умножения операндов с разрядностью выше 18.

Для реализации варианта с последовательным поступлением данных на вход (рис. 3) потребуется: $Q_{max} \times 3 \times 2 = 11 \times 3 \times 2 = 66$ DSP-блоков.

В табл. 1 приведены результаты реализации второго варианта построения формирователя лучей (рис. 3) на базе ПЛИС 10CL120YF484I7, полученные из отчетов САПР Quartus.

Объем ресурсов ПЛИС, затраченных на реализацию формирователя лучей

Ресурсы	Автоматическое каскадирование	Принудительное каскадирование (без IP-ядер)	Принудительное каскадирование (с IP-ядрами)
Число логических ячеек (Logic Cells)	9422 (8%)	9792 (8%)	8618 (7%)
Число регистров (Dedicated Logic Registers)	6463 (6%)	6562 (6%)	6253 (5%)
Объем встроенной памяти, бит (Memory Bits)	9004 (<1%)	9004 (<1%)	9004 (<1%)
Число блоков памяти (M9Ks)	12 (3%)	12 (3%)	12 (3%)
Число DSP-блоков (18×18)	66 (35%)	66 (35%)	66 (35%)
Максимальная тактовая частота ($f_{clk,max}$)	135,94 МГц	144,34 МГц	163,2 МГц

Здесь рассмотрены три варианта реализации процедуры каскадирования умножителей, проиллюстрированной на рис. 4:

- автоматическая реализация средствами САПР Quartus;
- принудительная реализация средствами языка описания аппаратуры без использования IP-ядер;
- принудительная реализация средствами языка описания аппаратуры с использованием IP-ядер.

Как видно, наименьший объем затрачиваемых ресурсов и наибольшую тактовую частоту работы схемы показывает вариант принудительного каскадирования с использованием IP-ядер. Причем этот вариант реализации формирователя лучей занимает небольшой процент от общего объема ресурсов, доступных в ПЛИС, и обеспечивает требование по минимальной тактовой частоте (не менее 144 МГц) с запасом около 20 МГц.

Заключение

Предложенные варианты реализации модуля цифрового диаграммообразования (формирователя лучей), основанные на применении последовательно-параллельного способа организации вычислений, позволяют обеспечить приемлемый уровень затрат вычислительных ресурсов и являются предпочтительными для использования в составе программного обеспечения ПЛИС, реализующего цифровую обработку сигналов автомобильного радара.

Проведенный анализ особенностей применения специализированных DSP-блоков, являющихся частью архитектуры ПЛИС семейства Cyclon 10LP фирмы Intel, для выполнения операций умножения операндов повышенной разрядности (более 18) позволяет сделать вывод о целесообразности использования IP-ядер умножителей, входящих в комплект поставки САПР Quartus, вместо реализации операции умножения средствами языка описания аппаратуры (HDL). Предложенный вариант оптимизации архитектуры комплексных умножителей позволяет снизить число используемых DSP-блоков на 25%.

Приведенные результаты реализации разработанного модуля в виде таблицы с объемом затраченных ресурсов ПЛИС разного типа могут быть полезны при планировании необходимого объема аппаратных ресурсов, требуемых для реализации алгоритмов цифровой обработки сигналов автомобильных радаров, с целью выбора подходящей модели микросхемы ПЛИС. Приведенная информация о максимальной частоте работы реализованной схемы позволяет оценить возможности ПЛИС семейства Cyclon 10LP в части быстродействия, а также может оказать помощь в планировании сетки тактовых частот при разработке проекта ПЛИС.

Библиографический список

1. Jonathan Bechter, Karim Eid, Fabian Roos, Christian Waldschmidt «Digital Beamforming to Mitigate Automotive Radar Interference» // Proceedings in 2016 IEEE MTT-S International Conference on Microwaves for Intelligent Mobility (ICMIM), 19-20.05.2016, USA
2. Intel corp. «Intel Cyclone 10 LP Device Datasheet»: [электронный ресурс]. URL: <https://www.intel.com/content/dam/www/programmable/us/en/pdfs/literature/hb/cyclone-10/c10lp-51002.pdf> (03.02.2021).
3. Intel corp. «Intel Cyclone 10 LP Device Design Guidelines»: [электронный ресурс]. URL:<https://www.intel.com/content/dam/www/programmable/us/en/pdfs/literature/an/an800.pdf> (03.02.2021).
4. Блейхут Р. Быстрые алгоритмы цифровой обработки сигналов. М.: «Мир», 1989, 447 с.

I.K. Malyshevsky, S.E. Kuznetsov

DEVELOPMENT AND OPTIMIZATION OF THE DIGITAL BEAMFORMING UNIT FOR AUTOMOTIVE RADAR

Nizhny Novgorod state technical university n.a. R.E. Alekseev

The issues of development of digital beamforming unit for automotive radar on FPGA base are considered, and the analysis of architecture optimization ways is carried out in order to reduce the hardware resources and improve the time characteristics.

Keywords: automotive radar, beamforming, FPGA, resource optimization

ПРИМЕНЕНИЕ ПРОСВЕТНОГО РАДАРА В СИСТЕМЕ БЕЗОПАСНОСТИ ЖЕЛЕЗНОДОРОЖНЫХ ПЕРЕЕЗДОВ

Нижегородский государственный технический университет им. Р.Е. Алексеева

Радиолокационное обнаружение и сопровождение малозаметных наземных целей представляет большой интерес для обеспечения безопасности движения на железнодорожных переездах. В статье рассмотрены принципы построения радиолокационного комплекса системы безопасности железнодорожного переезда, включая радар с обнаружением «на просвет». Представлены основные аспекты использования просветного радара и его структурная схема.

Ключевые слова: просветная РЛС; система безопасности железнодорожного переезда; диаграмма направленности; обнаружение неподвижных объектов.

Введение

Для обнаружения объектов на закрытых для автотранспорта железнодорожных (ЖД) переездах могут использоваться сенсоры, работающие в оптическом или радиочастотном диапазонах. К оптическим средствам относятся видеокамеры и лидары. Их основным недостатком является значительное ухудшение эффективности в условиях недостаточной видимости – при дожде, снегопаде, тумане.

Радарные системы в значительной степени лишены этого недостатка и могут рассматриваться как основные всепогодные средства обнаружения. Наиболее широко используются РЛС, работающие в диапазоне 24 ГГц. Причиной широкого распространения таких радаров является наличие доступной элементной базы, сравнительно низкая стоимость.

Основным недостатком радара, реализующего обнаружение сигнала, отраженного от цели в обратном направлении, является существенная зависимость эффективности обнаружения от отражательной способности обнаруживаемых объектов, которая зависит от материала, из которого изготовлен объект, а также от формы объекта, и численно характеризуется эффективной площадью рассеяния (ЭПР).

Дополнительно ухудшить условия обнаружения может также интерференция отраженного от цели сигнала с сигналом, переотраженным от локальных предметов в зоне переезда или от подстилающей поверхности. Стоит отметить, что «рельеф» вокруг переезда может заметно меняться в зимний период в зависимости от количества выпавшего снега, который, при этом, обладает хорошими отражающими свойствами. Таким образом, использование радаров, реализующих обнаружение отраженного «назад» сигнала, в сложных метеоусловиях не позволяет обеспечить требуемый уровень безопасности на переезде.

Для повышения уровня безопасности (эффективности обнаружения целей) ЖД переездов помимо моностатических радаров предлагается использовать радар прямого рассеяния (РПР) или радар, работающий «на просвет». Статья посвящена анализу потенциальных характеристик и параметров РПР в составе РЛС ЖД переезда.

Достоинства и недостатки классической схемы просветной РЛС с внешней когерентностью

ЭПР объектов при обнаружении «на просвет» определяется только теневым силуэтом и не зависит ни от коэффициента отражения материала, ни от формы объекта [1].

Кроме того, ЭПР при обнаружении «на просвет» намного выше, чем при обнаружении рассеянного назад поля, если размеры объекта значительно превышают длину волны излучения.

Узость зоны обнаружения просветного радара является платой за повышение просветной ЭПР. При рассеянии вперед вся энергия рассеянного сигнала концентрируется в окрестности направления, соответствующего направлению облучения, и быстро падает при увеличении угла между направлением облучения и направлением рассеяния (угла дифракции). В [2] был предложен критерий выбора оптимального частотного диапазона для построения просветного радара при обнаружении объектов заданного класса. Кроме того, для увеличения покрытия обычно используется мультистатическая конфигурация РПР [3, 4].

С другой стороны, для мониторинга ЖД переезда очень важным аспектом является возможность обнаружения стационарных объектов в зоне переезда. Именно остановка транспорта на путях является наиболее опасной ситуацией. Ограничения по возможности обнаружения стационарных объектов при использовании просветных радаров обусловлены тем, что при нахождении объекта в окрестности линии базы для разрешения отраженного сигнала и прямого сигнала передатчика требуется весьма широкая полоса сигнала. Используя соотношения, приведенные в [3], можно показать, что при длине линии базы 50 м для обнаружения неподвижного объекта, расположенного на расстоянии 5 м от линии базы, необходима ширина спектра не менее 250 МГц.

При использовании узкополосного (квазигармонического) непрерывного зондирующего сигнала просветный радар обеспечивает измерение только сдвига частоты отраженного сигнала и угла прихода, и не позволяет определить координаты цели за один обзор. Даже для предварительной (первичной) оценки координат целей необходимо использовать несколько обзоров [5]. В рамках радиочастотного плана, действующего в РФ, полосу частот шириной 83 МГц, можно получить в диапазоне 2,4...2,483 ГГц, а полосу 150 МГц – в диапазоне 5,125...5,275 ГГц.

На рис. 1 показаны элементы разрешения просветной РЛС по суммарной задержке, находящиеся между границами эллипсов равных суммарных дальностей, для двух значений ширины спектра сигнала – 80 МГц и 150 МГц. Значительная часть зоны обнаружения является наибольшей центральной областью разрешения, представляющей собой эллипс с малой полуосью (расстоянием от линии базы до границы) от 7 до 10 м.

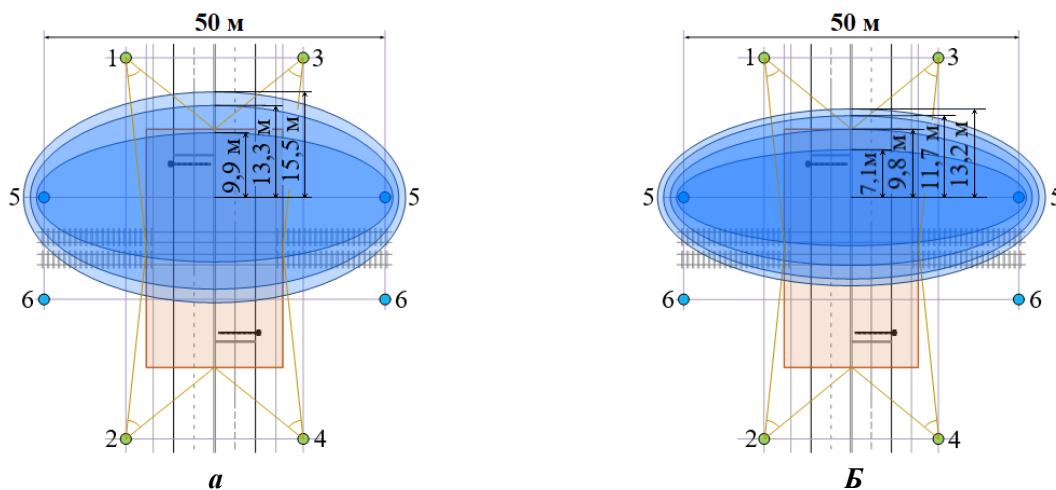


Рис. 1. Элементы разрешения по суммарной дальности:
1-4 – моностатические РЛС; 5, 6 – бистатические просветные РЛС;
а) для ширины спектра сигнала 80 МГц; б) для ширины спектра сигнала 150 МГц

Модифицированная схема построения просветной РЛС

Классическая схема построения просветного радара предполагает излучение сигнала передатчика в направлении приемной позиции и работу приемной позиции в режиме внешней когерентности [1].

Таким образом, в центральном элементе разрешения отраженный от неподвижного объекта сигнал практически не может быть обнаружен, поскольку он не разрешается на фоне прямого сигнала.

Проблема обнаружения неподвижного объекта в центральном элементе разрешения (рис.1) может быть решена за счет исключения (или многократного снижения) воздействия прямого сигнала передатчика на тракт приемной позиции путем формирования нулей ДН передающей и приемной антенн в направлении вдоль линии базы. Лучи ДН передатчика и приемника могут быть сформированы в направлении центра зоны переезда, как это показано на рис.2. За счет такого углового смещения лучей ДН обеспечивается увеличение коэффициента усиления антенн при облучении цели и приеме рассеянного сигнала в зоне ЖД переезда.

Поскольку классическая для просветного радара внешняя синхронизация недоступна из-за подавления прямого сигнала, требуется обеспечить наличие когерентного опорного сигнала на приемной стороне альтернативным способом. Когерентный опорный сигнал может передаваться по коаксиальному кабелю, что вполне приемлемо при рассматриваемых расстояниях между позициями передачи и приема (50...70 м). Таким образом, при отсутствии прямого сигнала передатчика в приемном тракте может быть реализовано обнаружение объекта, находящегося внутри указанного центрального элемента разрешения.

В качестве зондирующих сигналов предлагается использовать непрерывные сигналы с линейной частотной модуляцией (СВФМ). В этом случае для обработки удобно использовать прямой перенос спектра на нулевую частоту (гомодинный прием) за счет умножения принимаемого сигнала на излучаемый [6, 7].

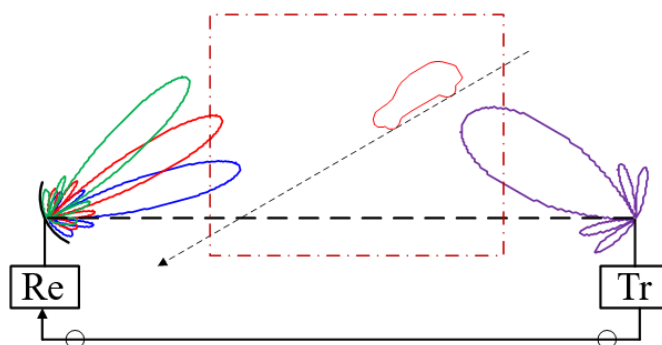


Рис. 2. Угловое смещение ДН антенн для снижения воздействия прямого сигнала передатчика на тракт приемной позиции

В качестве антенн в рассматриваемых частотных диапазонах предлагается использовать плоские печатные фазированные антенные решетки (ФАР). На рис. 3, а представлен внешний вид передающей и приемной антенн для диапазона 2,4 ГГц. ФАР представляют собой набор столбцов из нескольких излучающих элементов – патчей, соединенных микрополосками. ДН передающей антенны – однолучевая, освещающая всю зону наблюдения. Лучи ДН приемной антенны формируются путем цифровой обработки.

На рис. 3, б показаны ДН передающей (кривая 1) и приемной (кривые 2-4) антенн. Полученные диаграммы имеют минимальный уровень вдоль линии базы (нулевое направление) и смещены к центру переезда. Для формирования нулей ДН в направлении вдоль линии базы и максимумов в заданных направлениях использован пространственный фильтр Винера [8].

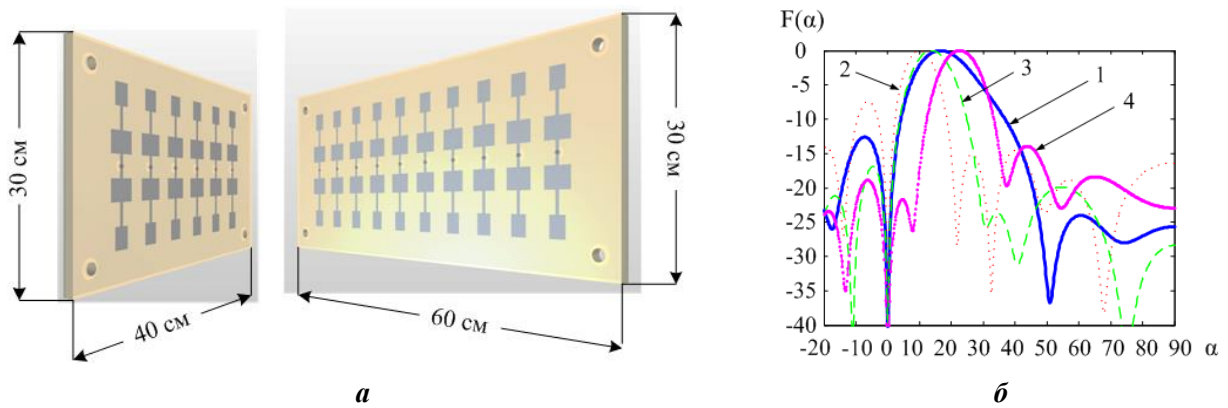


Рис. 3. а) Передающая и приемная антенные решетки для РПР диапазона 2,4 ГГц; б) ДН передающей (кривая 1) и приемной (кривые 2-4) антенн

На рис. 4 показана совместная структурная схема аналоговой части, передающей и приемной позиций. Непрерывный сигнал с периодической кусочно-линейной частотной модуляцией (CWLFM) формируется генератором, управляемым напряжением (ГУН). Параметры формируемого сигнала (девиация частоты, период перестройки частоты) определяются цифровым блоком управления (ЦБУ). После ограничения полосы сигнала полосовым фильтром ПФ1 и усилителя мощности УМ2 сигнал поступает на вход согласующей линии передающей антенны. Отраженный от объектов сигнал с выхода приемной антенны после малошумящего усилителя (МШУ) и ПФ2 поступает на вход смесителя (СМ), в котором перемножается с усиленным УМ1 опорным сигналом, который с выхода ГУН передающей позиции передается по коаксиальному кабелю. На выходе СМ сигнал разностной частоты выделяется фильтром нижних частот (ФНЧ). Затем паразитная составляющая на низкой частоте, обусловленная попаданием на вход смесителя прямого сигнала, ослабляется при помощи фильтра верхних частот (ФВЧ). После ФВЧ сигнал проходит через усилитель низких частот (УНЧ) и поступает на вход аналого-цифрового преобразователя (АЦП), с выхода которого цифровые отсчеты поступают в систему цифровой обработки сигналов (ЦОС).

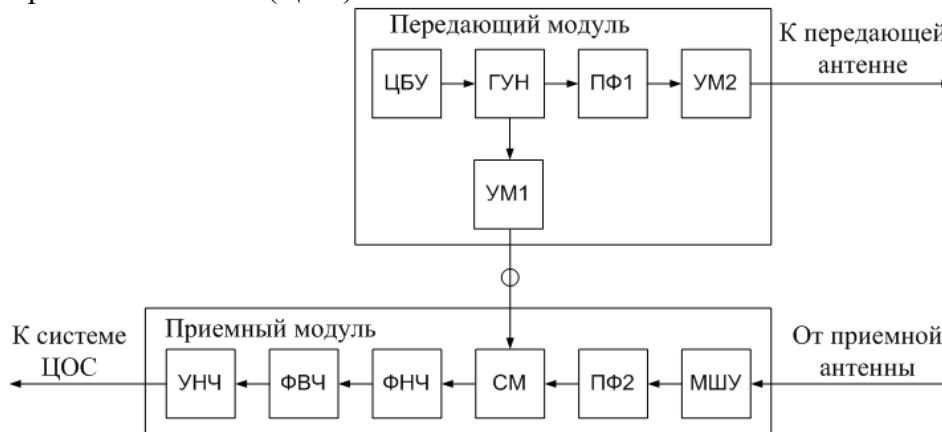


Рис. 4. Структурная схема передающего и приемного модулей

Недостатком предлагаемой методики является то, что зона действия «смещенных» ДН антенны имеет достаточно широкий зазор вдоль линии базы. Таким образом, цели, приближающиеся к ЖД переезду, не обнаруживаются. Для решения данной проблемы предлагается двухдиапазонное исполнение. Просветный радар, работающий в диапазоне 2,4...2,483 ГГц подавленным прямым сигналом, дополняется РПР, работающим в диапазоне 5,125...5,275 ГГц. «Высокочастотный» радар, размещается на тех же позициях и работает в соответствии с классической схемой построения просветного радара в режиме внешней когерентности, полученной за счет наличия прямого сигнала на приемной стороне.

Анализ эффективности обнаружения целей

На рис. 5 показаны зоны обнаружения просветным радаром куба с размерами $0,5 \times 0,5 \times 0,5$ м (рис. 5,а) и автомобиля (рис. 5,б), сформированные с учетом описанных направленных свойств антенн. Синим цветом показана ЗО радара, работающего в диапазоне 2,4...2,483 ГГц, бордовым цветом – ЗО радара, работающего в диапазоне 5,125...5,275 ГГц. ЗО построены таким образом, что внутри этих областей отношение сигнал-шум составляет не менее 42 дБ, что обеспечивает обнаружение за один обзор (0,1 с) с вероятностью 0,999 при вероятности ложной тревоги 10^{-6} .

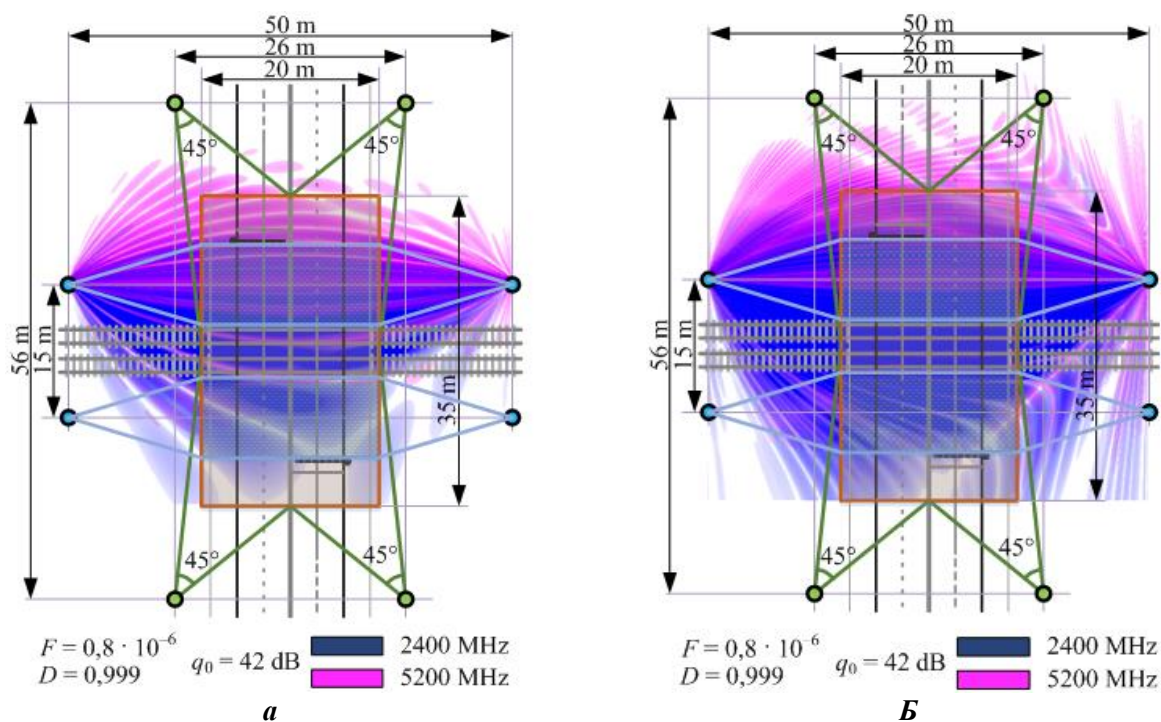


Рис. 5. Зоны обнаружения: а) куба размером $0,5 \times 0,5 \times 0,5$ м; б) автомобиля

Результаты и выводы

В работе рассмотрены принципы построения радиолокационного комплекса для системы безопасности ЖД переезда. В частности, предложена реализация входящей в его состав просветной РЛС.

Выработаны ключевые аспекты построения просветного радара. Для обеспечения возможности обнаружения неподвижных объектов необходимо исключить наличие прямого сигнала передатчика в приемном канале; обеспечить прямое измерение суммарной дальности за счет использования сигнала с относительно широкой полосой (доступные полосы частот 83 МГц или 150 МГц).

Наиболее подходящей конструкцией антенны в данном случае является ФАР на микрополосковых отрезках (патч-антенна) с горизонтальным размером передающего модуля около 35 см и приемного модуля около 60 см.

Опорный когерентный сигнал синхронизации на приемную позицию передается по кабелю. Длина линии базы 50...70 м позволяет проложить коаксиальный кабель, при передаче по которому будет обеспечен приемлемый уровень потерь.

При подавлении прямого сигнала передатчика сигнал, рассеянный неподвижным объектом, может быть обнаружен уже в первом элементе разрешения – центральном эллипсе равных суммарных дальностей (рис. 1).

Для обеспечения непрерывной зоны обнаружения и расширения ее в противоположную от переезда сторону относительно линии базы целесообразно использовать двухдиапазонный комплекс. При этом на второй (более высокой) частоте реализуется стандартный принцип построения с излучением сигнала вдоль линии базы.

Библиографический список

1. Cherniakov M. Bistatic radar: principles and practice. Wiley. 2007. – 504 p.
2. Blyakhman A.B., Burov V.N., Myakinkov A.V., Ryndyk A.G. Detection of Unmanned Aerial Vehicles via Multi-Static Forward Scattering Radar with Airborne Transmit Positions // Proc. of International Radar Conference. Lille, France. 13-17 October 2014.
3. Мякинков А.В., Смирнова Д.М. Обнаружение наземных целей в многопозиционной просветной радиолокационной системе // Известия высших учебных заведений России. Радиоэлектроника. 2010. № 5. С. 47-55.
4. Daniel L.Y., Gashinova M.S., Cherniakov M., Target Observability Improvement in Multi-Static Maritime Forward Scatter Radar // Proc. of International Radar Conference. Lille, France. 2014.– 6 p.
5. Бляхман А.Б., Мякинков А.В., Рындык А.Г. Измерение координат целей в трехкоординатных бистатических РЛС с обнаружением «на просвет» // Радиотехника и электроника. 2006. №4. С.422-427.
7. Rohling H. Automotive radar// Proc. SPIE 5484. Photonics Applications in Astronomy, Communications, Industry, and High-Energy Physics Experiments II. 22 July 2004. doi: 10.1117/12.569057.
8. Rohling H., Meinecke M.-M. Waveform Design Principles for Automotive Radar System// CIE International Conference on Radar. 2001.
9. Monzingo R.A., Haupt R.L., Miller T.W. Introduction to Adaptive Arrays. 2nd edition. SciTech Publishing. 2011.– 544 p. ISBN: 978-18911215.

A.V. Myakinkov, K.S. Fomina, S.A. Shabalin, A.D. Mikhaylov

RAILWAY CROSSINGS SAFETY RADAR SYSTEM WITH USE OF FORWARD SCATTER RADAR

Nizhny Novgorod state technical university n.a. R.E. Alekseev

Radar detection and tracking of low-observable ground targets is of great interest to ensure the safety of traffic at railway crossings. The paper deals with the building a radar complex with forward scatter radar for a railway crossing safety system features. The specific aspects of the forward scatter radar use and its structural scheme are demonstrated.

Keywords: forward scatter radar; railway crossing safety system; radiation pattern, detection of stationary objects.

АЛГОРИТМ ВЫСОКОТОЧНОГО ОЦЕНИВАНИЯ ПЕРЕМЕЩЕНИЯ АВТОМОБИЛЯ ДЛЯ РЕШЕНИЯ ЗАДАЧИ ПОВЫШЕНИЯ УГЛОВОГО РАЗРЕШЕНИЯ АВТОМОБИЛЬНОГО РАДАРА

Нижегородский государственный технический университет им. Р.Е. Алексеева

Разработан алгоритм оценки перемещения автомобиля для построения синтезированной апертуры автомобильным радаром миллиметрового диапазона.

Ключевые слова: автомобильный радар, синтез апертуры, оценка перемещения, акселерометр, интегрирование.

Введение

Разработка и развитие систем интеллектуальной помощи водителю на основе автомобильных радиолокаторов неразрывно связаны с необходимостью повышения разрешающей способности по всем координатам. Повышение разрешающей способности по угловым координатам остаётся сложной задачей, требующей новых эффективных решений. Одним из возможных способов решения является применение алгоритмов синтеза апертуры. Синтез апертуры в классическом виде пригоден для сканирования в пределах боковых секторов обзора автомобиля. При сокращении угла между сканирующим лучом и вектором перемещения антенны, эффективность данного алгоритма резко уменьшается. Этот факт существенно ограничивает возможности применения классического синтеза апертуры в переднем и заднем секторах обзора.

Некоторые способы решения указанной проблемы описаны в [1,2]. Все эти способы применяются для адаптации синтеза апертуры к работе в переднем секторе обзора при равномерном прямолинейном движении автомобиля. На практике можно воспользоваться тем фактом, что водители часто вынуждены корректировать направление движения, что делает реальную траекторию движения зигзагообразной. Таким образом, можно выделить усреднённую «правильную» траекторию движения, относительно которой присутствует поперечная составляющая вектора перемещения. Это поперечное перемещение можно использовать для синтеза апертуры в переднем секторе обзора автомобиля.

Использование зондирующего сигнала миллиметрового диапазона означает необходимость измерения траектории движения антенны с точностью до долей миллиметра [3]. Соответственно, вычислитель должен учитывать различные флуктуации автомобиля, которые невозможно вычислить по данным приборной скорости и положению рулевого колеса из-за отработки шинами и подвеской различных неровностей, а также воздействия прочих сил, например, бокового ветра. В таких условиях единственным способом оценки траектории с заданной точностью является использование акселерометра с соответствующим алгоритмом обработки его данных.

В данной статье приводится обзор и результаты первого шага разработки алгоритма оценки поперечного перемещения для автомобильной РЛС, заключающегося в оценке с заданной точностью линейного перемещения объекта на основании данных акселерометра.

Постановка задачи

Синтез апертуры заключается в когерентной обработке сигналов, принятых движущейся антенной в разных пространственных, а как следствие и временных её положениях за фиксированный интервал времени. Точность оценки траектории движения антенны непосредственно влияет на качество полученного радиоизображения местности.

Для синтеза апертуры важны не столько реальные координаты антенны, сколько оценка её относительного перемещения за интервал синтеза.

Входными данными для алгоритма оценки являются данные с акселерометра. Физически эти данные представляют собой напряжение в пределах некоторого диапазона. Правила соответствия напряжения ускорению зависят от конкретной микросхемы акселерометра, а именно: от чувствительности (мВ/г), от диапазона измеряемых ускорений. Поэтому, в рамках данного исследования, будем работать непосредственно с ускорением, которое можно представить следующим образом:

$$\begin{aligned} a(t) &= a_{\text{действ.}}(t) + n(t), \\ n(t) &= n_{\text{кв}}(t) + n_{\text{ш}}(t), \end{aligned} \quad (1)$$

где $a(t)$ – измеренное ускорение, $a_{\text{действ.}}(t)$ – действительное ускорение антенны, $n(t)$ – шумовая составляющая принятых в обработку данных, $n_{\text{кв}}(t)$ – шум квантования, $n_{\text{ш}}(t)$ – прочие шумы.

Сравнивая действительную и измеренную траектории в пределах интервала синтеза, в качестве критерия оценки их взаимного соответствия выделим максимальную флуктуационную ошибку измеренной траектории за период синтеза.

Теоретическое обоснование

Связь между функциями ускорения, скорости и перемещения имеет вид:

$$\begin{aligned} x(t) &= \int v(t) dt, \\ v(t) &= \int a(t) dt = \int (a_{\text{действ.}}(t) + n(t)) dt = \\ &= \int a_{\text{действ.}}(t) dt + \int n(t) dt. \end{aligned} \quad (2)$$

Из выражения (2) видно, что такой подход позволяет эффективно подавлять шум. В соответствии с теоремой о среднем:

$$\lim_{T \rightarrow \infty} \int n(t) dt = M\{n(t)\}. \quad (3)$$

где $M\{n(t)\}$ – математическое ожидание случайного процесса $n(t)$. Таким образом, с течением времени ошибка оценки перемещения будет стремиться к минимально возможному значению.

Оценка перемещения ведётся в пределах интервала синтеза T_c , задача должна решаться в определённых интегралах вида:

$$\begin{aligned} x(t) &= \int_{t-T_c}^t v(t) dt + C_1, \\ v(t) &= \int_{t-T_c}^t a(t) dt + C_2, \\ x(t) &= \int_{t-T_c}^t (\int_{t-T_c}^t a(t) dt + C_2) dt + C_1 =, \\ &= \int_{t-T_c}^t \int_{t-T_c}^t a(t) dt dt + C_2 T_c + C_1. \end{aligned} \quad (4)$$

Из выражений (4) ясно, что при определённом интегрировании ускорения и скорости на интервале синтеза, возникает ошибка оценки траектории, из-за появления констант C_1 и C_2 . Эти константы фактически представляют собой соответственно координату и скорость антенны на момент начала интервала синтеза, то есть:

$$\begin{aligned} C_1 &= x(t - T_c) = x_0, \\ C_2 &= v(t - T_c) = v_0. \end{aligned} \quad (5)$$

Константа C_1 , определяя лишь исходное положение антенны, никак не влияет на форму траектории, а значит никак не влияет на точность оценки относительного положения различных точек траектории, следовательно, константой C_1 можно пренебречь. Константа C_2 напротив, влияет на флуктуационную ошибку оценки траектории. Чем больше значение C_2 будет отличаться от нуля и чем больше будет период синтеза T_c , тем меньше форма оцененной траектории будет соответствовать форме реальной траектории. Учесть константу C_2 в вычислениях возможно путём интегрирования функции ускорения $a(t)$ до текущего времени t от некоторого момента с известной начальной скоростью. Учитывая предположение о том, что вычислитель начинает работу с момента запуска двигателя на неподвижном автомобиле, таким моментом может быть самое начало работы вычислителя, когда $C_2 = 0$. Таким образом, интегрирование ведётся непрерывно на протяжении всего времени поездки. При этом, в памяти вычислителя должны храниться лишь отсчёты в пределах текущего интервала синтеза T_c и начальные условия для того же интервала.

Построение алгоритма

Исходя из приведённых выше рассуждений, становится ясен вид алгоритма получения оценки перемещения: он состоит из двух интеграторов, осуществляющих расчёт функции перемещения на интервале синтеза T_c . Результаты расчёта обновляются каждый раз при поступлении на вычислитель нового отсчёта ускорения; функция перемещения пересчитывается на интервале синтеза T_c , сдвинутым на интервал дискретизации T_d . Два интегратора отличаются друг от друга за счёт необходимости учёта константы интегрирования в одном случае и отсутствия такой необходимости в другом.

В обоих случаях интегрирование будем производить методом трапеций (рис.1). Такой метод позволяет существенно увеличить точность интегрирования без использования сложных математических операций.

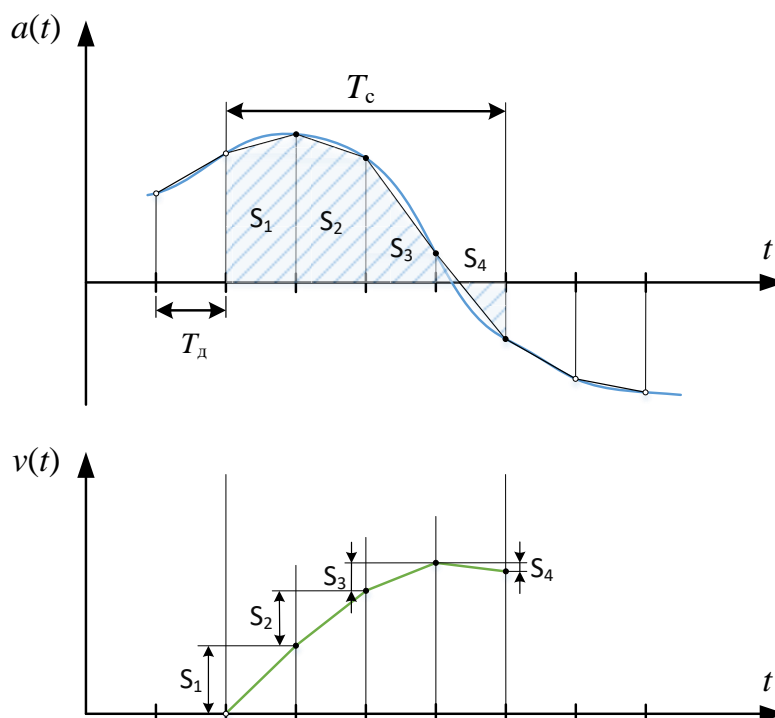


Рис. 1. Интегрирование на интервале

Второй интегратор выполняет классическую задачу нахождения первообразной на определённом интервале. Его структурная схема представлена на рис. 2.

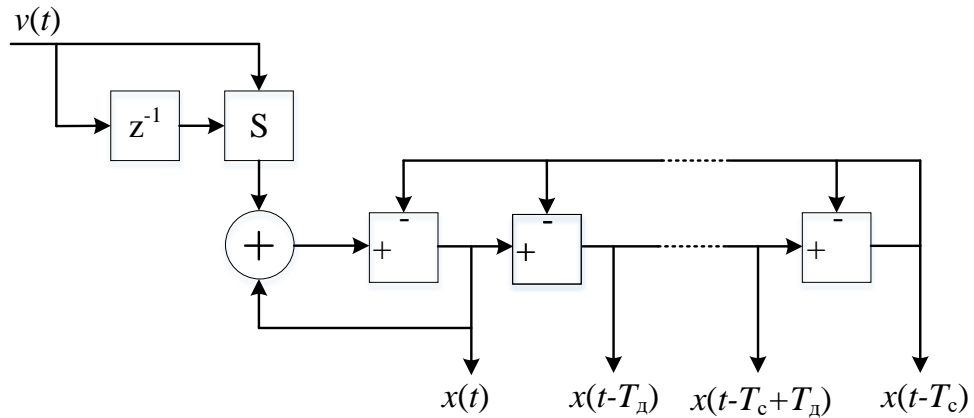


Рис. 2. Структурная схема второго интегратора

На вход вычислителя последовательно поступают отсчёты скорости. Блок S на основании текущего и предыдущего отсчёта осуществляет расчёт площади под функцией на интервале T_d , используя метод трапеций.

$$S_v(n) = v(n-1)T_d + \frac{(v(n)-v(n-1))T_d}{2}. \quad (6)$$

Значения площадей поступают на накопитель, выход которого соединён со входом регистра, образованного синхронными устройствами вычитания. С каждым тактом к предыдущему значению отсчёта $x(t)$ прибавляется значение площади под функцией на следующем интервале T_d . Одновременно с этим, из каждого выходного отсчёта вычитается последнее значение $x(t-T_c)$. Таким образом, отсчёты функции перемещения на интервале синтеза T_c представляют собой суммы всех площадей под функцией скорости с момента начала интервала синтеза до момента времени, соответствующего данному отсчёту.

Первый интегратор должен производить расчёт функции скорости не только на основании попадающих в интервал синтеза отсчётов ускорения, но и всех предыдущих. Такой интегратор может быть получен путём модифицирования второго интегратора, рассмотренного выше. Если удалить из него канал вычитания ушедших из регистра отсчётов, то значения, хранящиеся в регистре, будут соответствовать первообразной функции ускорения, рассчитанной от момента начала процесса измерения (рис. 3). Накопитель в данном случае выполнен в виде синхронного сумматора.

Удаление канала вычитания, очевидно, упрощает схему, при этом позволяет учесть возникающую при интегрировании константу. Однако на втором интеграторе такое упрощение невозможно ввиду того, что разрядная сетка рано или поздно переполнится. Первый интегратор, в отличие от второго, защищён от переполнения техническими характеристиками самого автомобиля: наличие максимальной скорости делает невозможным её перманентный рост.

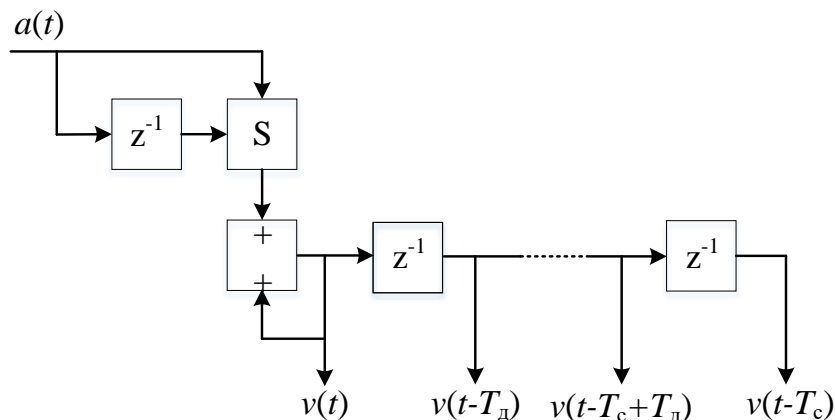


Рис. 3. Структурная схема первого интегратора

Объединив интеграторы, получаем структурную схему вычислителя, изображённую на рис. 4. При объединении двух интеграторов заметим, что оба интегратора рассчитаны на обработку последовательно поступающих на вход отсчётов исходной функции. Это означает отсутствие необходимости в сохранении всех отсчётов функции скорости на интервале синтеза, что существенно сокращает число структурных элементов схемы.

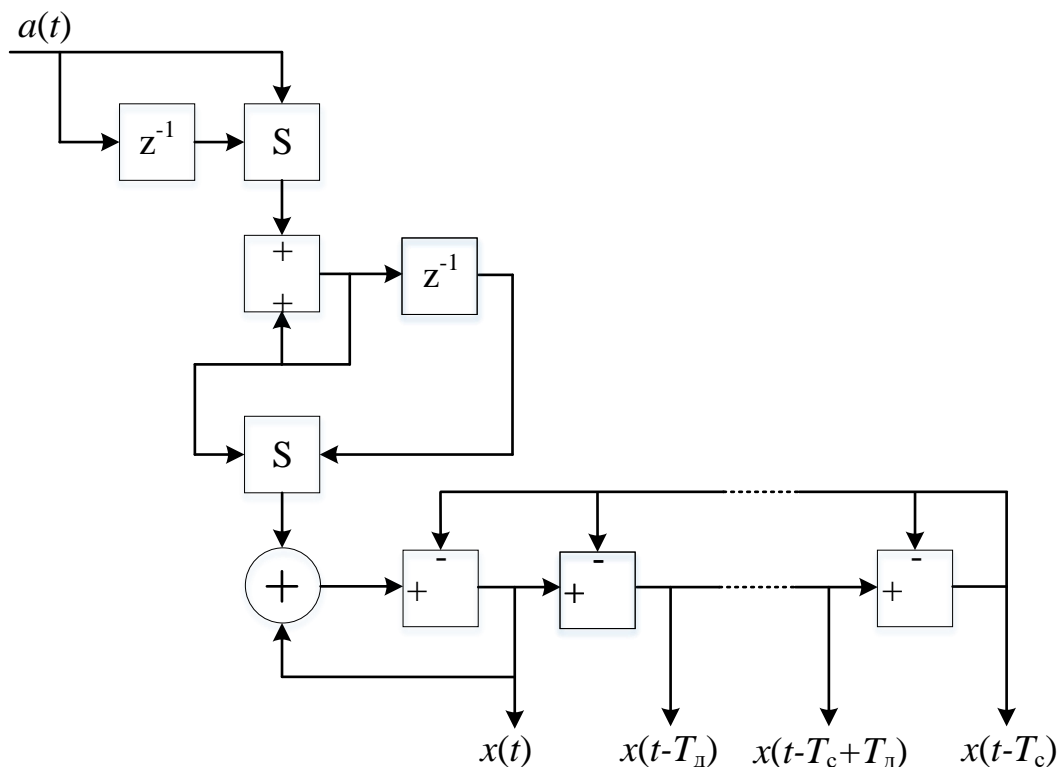


Рис. 4. Структурная схема алгоритма

Результаты моделирования

Компьютерное моделирование представленного алгоритма выполнено с помощью программы MATLAB. К синтезированной функции действительного ускорения примешивается шумовой сигнал с нормальным распределением, после чего подвергается искажению с помощью алгоритма имитации квантования. Отсчёты квантованной зашумлённой функции последовательно поступают на вход алгоритма. Результатом работы алгоритма является набор функций перемещения на сдвигаемых на величину T_d интервалах синтеза T_c .

Каждая из этих функций сравнивается с соответствующим ей по времени участком эталонной функции перемещения, получаемой из неискажённой функции действительного ускорения. Результатами сравнения являются значения максимальной флуктуационной ошибки на соответствующих интервалах. Данные значения объединяются в графики.

При моделировании алгоритма, в качестве опорного варианта акселерометра выберем микросхему ADXL326 фирмы Analog Devices. Максимальное значение измеряемых ускорений составляет до 16g по всем осям, чувствительность составляет 0,06 В/г.

Параметры квантования выбраны следующими: диапазон значений $-8g \dots +8g$ [4], 24 бит. Интервал синтеза $T_c = 0,5$ с, период дискретизации $T_d = 10^{-5}$ с. Исходя из параметров чувствительности выбранной микросхемы, положим мощность шума в пересчёте на измеренное ускорение равной $0,1$ м/с². На рис. 5 представлен график максимальной флуктуационной ошибки при постоянном ускорении $a = 1$.

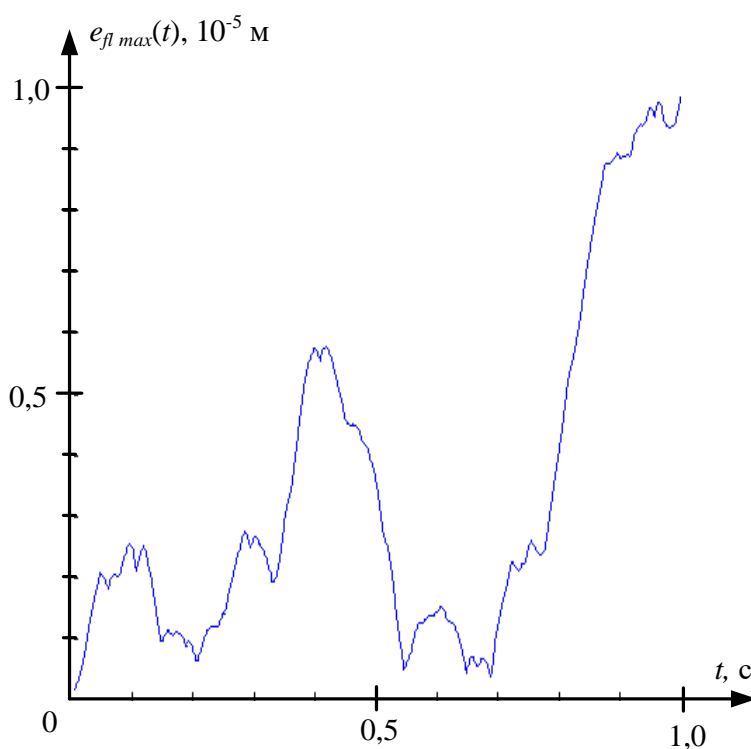


Рис. 5. Максимальная флуктуационная ошибка при постоянном ускорении

Из графика видно, что значение флуктуационной ошибки не превышает 0,01 мм. Рассмотрим поведение алгоритма при нелинейном ускорении (рис.6,7).

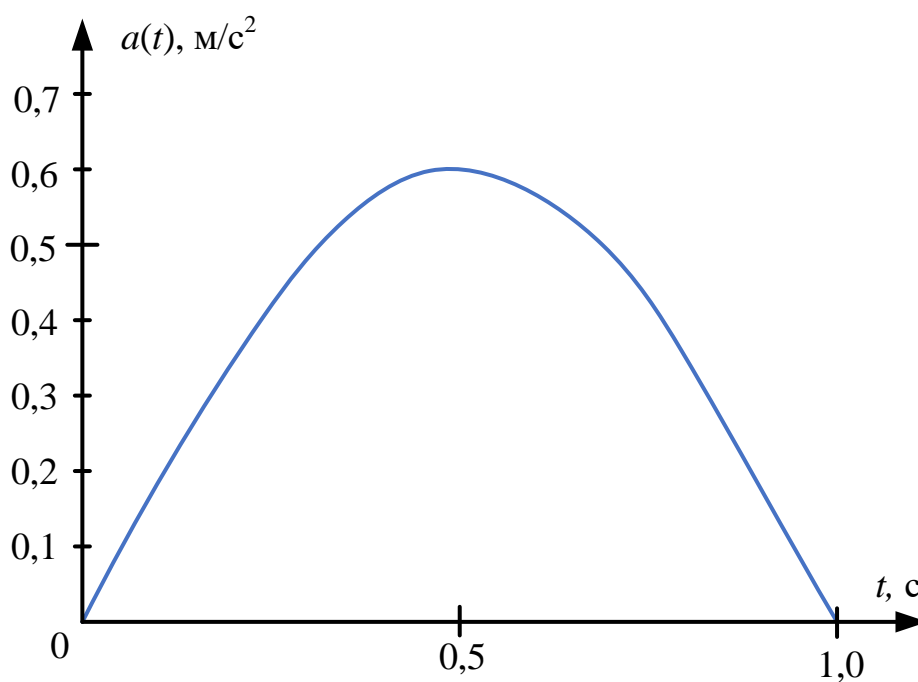


Рис. 6. График функции ускорения

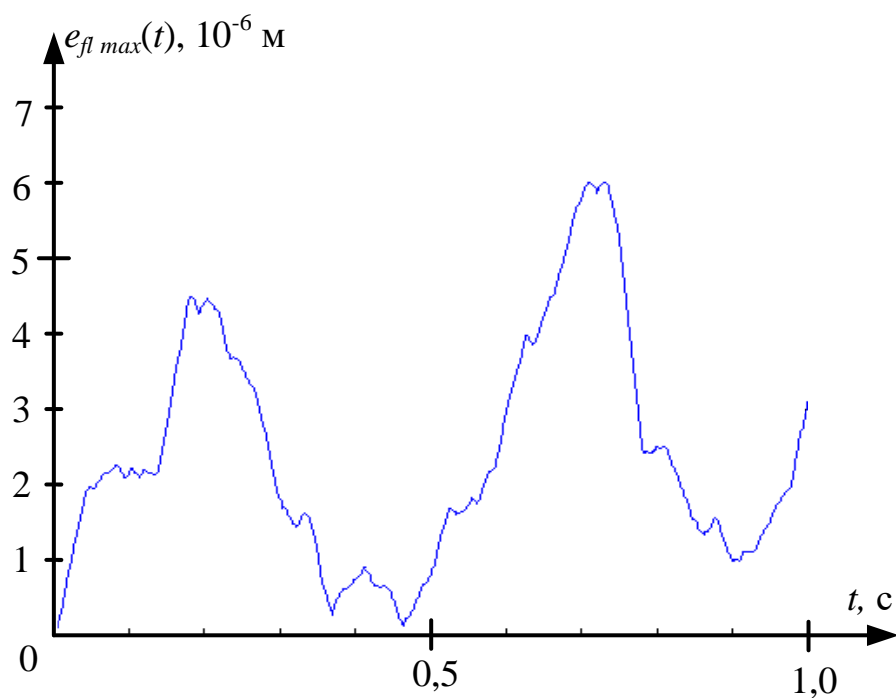


Рис. 7. Максимальная флуктуационная ошибка при нелинейном ускорении

При непостоянном и нелинейном ускорении алгоритм выдаёт результаты, схожие по точности с предыдущим случаем. Для того, чтобы убедиться в том, что ошибка не имеет тенденции к накоплению с течением времени, требуется промоделировать работу алгоритма на более длинном промежутке времени. На рис. 8 представлен график максимальной флуктуационной ошибки при работе алгоритма в течении 30 с. Ускорение в этом случае задано периодическим.

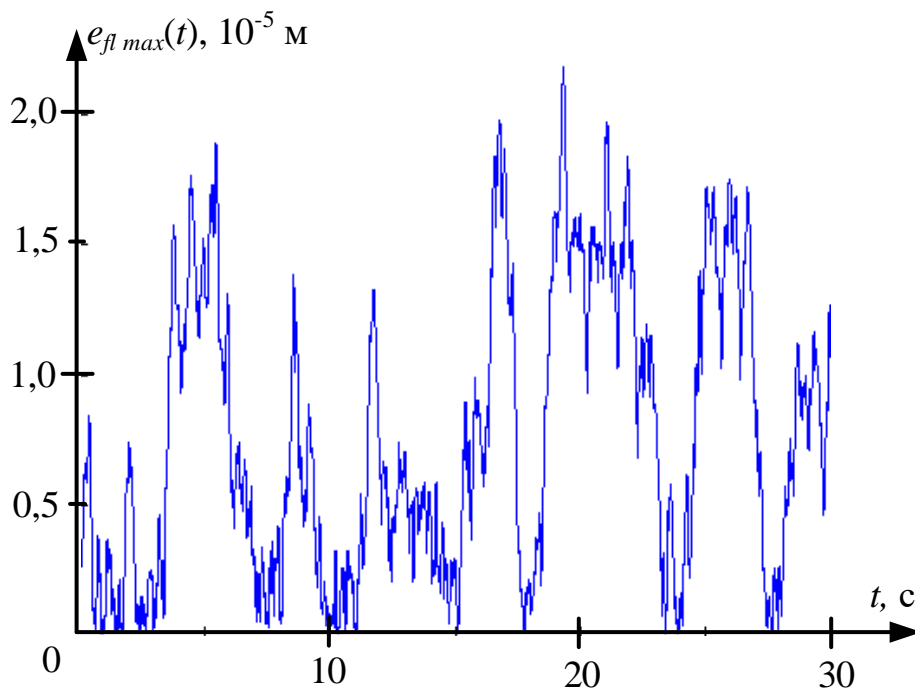


Рис. 8. Максимальная флуктуационная ошибка при нелинейном ускорении за 30с

Как видно из графиков максимальной флуктуационной ошибки, во всех представленных случаях алгоритм позволяет оценивать перемещение антенны с точностью, позволяющей обеспечивать когерентную обработку сигналов при синтезе апертуры.

Заключение

Представленное исследование позволяет сделать вывод о том, что выполнить оценку перемещения антенны по данным акселерометра при текущем развитии технологий и элементной базы возможно с точностью порядка сотых долей миллиметра. Это открывает возможности осуществления сложной пространственно-временной обработки в режиме реального времени и быстрого принятия решений, что необходимо для создания автоматической системы помощи водителю на основе радиолокационных устройств.

На основе исследования разработан алгоритм, осуществляющий оценку траектории перемещения во времени в пределах одного измерения. Данный алгоритм станет основой для дальнейшего расширения перечня оцениваемых параметров. Так, следующими этапами исследования являются: добавление возможности оценки перемещения во времени по нескольким пространственным координатам; развитие алгоритма для учёта специфики движения автомобиля, то есть учёт перемещения антенны с поворотом, для чего требуется разделить нормальную и тангенциальную составляющие ускорения из потока данных акселерометра по поперечной оси.

Библиографический список

1. Gishkori S., Daniel L., Gashinova M., Mulgrew B. Imaging for a Forward Scanning Automotive Synthetic Aperture Radar/ S. Gishkori, L. Daniel, M. Gashinova, B. Mulgrew// IEEE Transactions on Aerospace and Electronic Systems. – 2018.
2. Михайлов А.Д., Мякинков А.В. Анализ методов формирования радиолокационного изображения для автомобильного радара// Электронный сборник трудов XXVI-ой Международной Научно-технической конференции «Информационные системы и технологии» ИСТ–2020, Нижний Новгород, Радиотехнические системы и устройства (Секция 7). – 28 апреля 2020. – С.1340–1344.
3. Кондратенков Г.С., Фролов А.Ю. Радиовидение. Радиолокационные системы дистанционного зондирования земли. –М.: «Радиотехника», 2005. – 368с.
4. Хусаинов А.Ш., Селифонов В.В. Теория автомобиля. – Ульяновск, УлГТУ, 2008. – 121с.

A. D. Mikhailov, A. V. Myakinkov

AN ALGORITHM FOR HIGH-PRECISION ESTIMATION OF VEHICLE MOVEMENT TO SOLVE THE PROBLEM OF INCREASING THE ANGULAR RESOLUTION OF A CAR RADAR

Nizhny Novgorod state technical university n.a. R.E. Alekseev

An algorithm for estimating the movement of a car for constructing a synthesized aperture with a millimeter-wave car radar has been developed.

Keywords: automotive radar, aperture synthesis, displacement estimation, accelerometer, integration

АЛГОРИТМЫ ПОДАВЛЕНИЯ НЕСИНХРОННЫХ ИМПУЛЬСНЫХ ПОМЕХ В АВТОМОБИЛЬНОМ РАДАРЕ

Нижегородский государственный технический университет им. Р.Е. Алексеева

Борьба с активными помехами является одной из важнейших проблем в радиолокации. Это важно так же и для автомобильных радаров, применяемых в системе Adaptive Driver Assistance Systems (ADAS). Они должны обеспечивать надежную работу в условиях мощных активных помех. Без подавления таких помех на входе БПФ существенно увеличивается уровень маскирующего шума, в результате чего уменьшается отношение сигнал-шум, что ухудшает точность оценивания характеристик сигнала. От этих характеристик зависит точность измерения дальности, скорости и угла. В данном докладе идёт речь об алгоритмах борьбы с несинхронными импульсными помехами. Были рассмотрены два алгоритма подавления помехи, приведены результаты моделирования в среде MATLAB.

Ключевые слова: автомобильный радар, несинхронная импульсная помеха, скользящий порог.

Введение

Современные автомобили все чаще оснащаются активными системами помощи водителю, способными в зависимости от окружающей автомобиль дорожной ситуации оказывать управляющее воздействие на системы управления двигателем, тормозами, рулевым управлением. Поэтому в литературе получило широкое распространение их название Adaptive Driver Assistance Systems (ADAS). Одним из основных сенсоров, входящих в ADAS и обеспечивающих решение задачи оценивания обстановки перед автомобилем являются радары.

Автомобильные радары подвержены воздействию активных помех. Такие помехи представляют особую опасность и могут серьезно повлиять на характеристики автомобильных радаров. Активные помехи в автомобильных радаров возникают от встречно движущихся автомобилей с такими же типами радаров. Попадание в полосу частот приемника сигнала от мешающего радара приводит к возникновению импульсных помех [1]. На рисунке 1 показан пример межрадарной помехи в FMCW радаре с быстрым ЛЧМ, где T_U – длительность нарастания ЛЧМ сигнала, T_D – длительность спада ЛЧМ сигнала, Δf – ширина полосы частот сигнала, f_B – частота биений, f_{LPF} – полоса пропускания фильтра нижних частот. Мощность таких помех на порядок превышает мощность полезного отраженного сигнала.

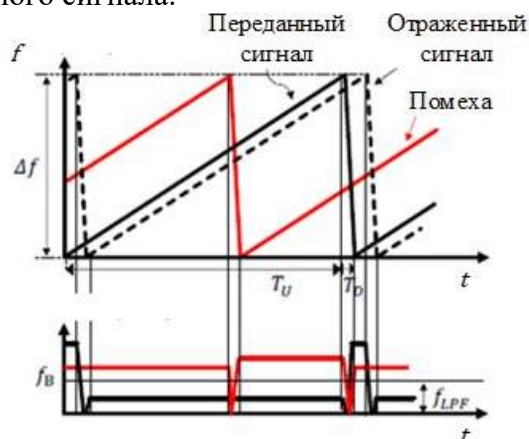


Рис. 1. Помехи в FMCW радаре

Алгоритмы компенсации помехи

Алгоритм компенсации помехи принято называть "ограничитель" амплитуды. Суть алгоритма состоит в том, чтобы как можно сильнее уменьшить амплитуду и, следовательно, мощность импульсной помехи и не изменить при этом амплитуду полезного сигнала. Рассмотрим два построения данного алгоритма: на основе скользящего окна и с применением постоянного порога, вычисляемого по всем точкам в периоде зондирования.

Значение порога для алгоритма с постоянным порогом вычисляется следующим образом:

$$h = \frac{1}{N} \sum_{n=0}^N s_n, \quad (1)$$

где s_n – модули отсчётов смеси сигнала с шумом, N – количество отсчетов в периоде зондирования.

На втором этапе данного алгоритма порог умножается на масштабирующий коэффициент, который подбирается экспериментально. Затем модуль каждого отсчёта сигнала сравнивается с этим числом. Если отсчет оказался больше значения порога по модулю, то считается, что в данном отсчете присутствует помеха и отсчёт нормируется к значению порога, если же отсчёт оказался меньше порога, то он остаётся без изменений.

Для алгоритма компенсации помехи на основе скользящего окна порог вычисляется по следующей формуле:

$$h(n) = \sum_{i=n-\frac{N}{2}}^{n+\frac{N}{2}} s_i \cdot W_i, \quad (2)$$

где W_i – весовая функция специальной формы.

Форма функции для данного окна будет описываться формулой:

$$W_i = 1 - \exp\left(-A \cdot \left(i - \frac{N}{2} + 0.5\right)\right), \quad (3)$$

где A – коэффициент, регулирующий крутизну формы.

Затем производится анализ сигнала в соответствии с вышеописанным алгоритмом.

Моделирование алгоритма постоянного порога в среде MATLAB

Для анализа подавления импульсных помех смоделируем один период зондирования. В модели сигнала зададим количество импульсных помех равное шестнадцати, действующих одновременно в один период зондирования. В результате сформируется сигнал заданной частоты со случайным временем действия импульсных помех (рис. 2)

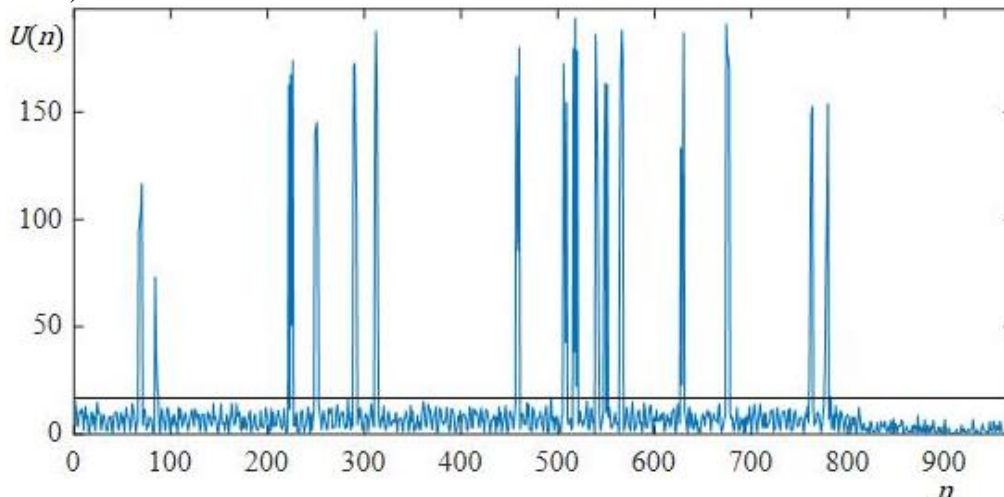


Рис. 2. Амплитуда анализируемого процесса с импульсной помехой и пороговый уровень

На рис. 3 представлена амплитуда процесса $U(n)$ после подавления помех при применении алгоритма с постоянным порогом.

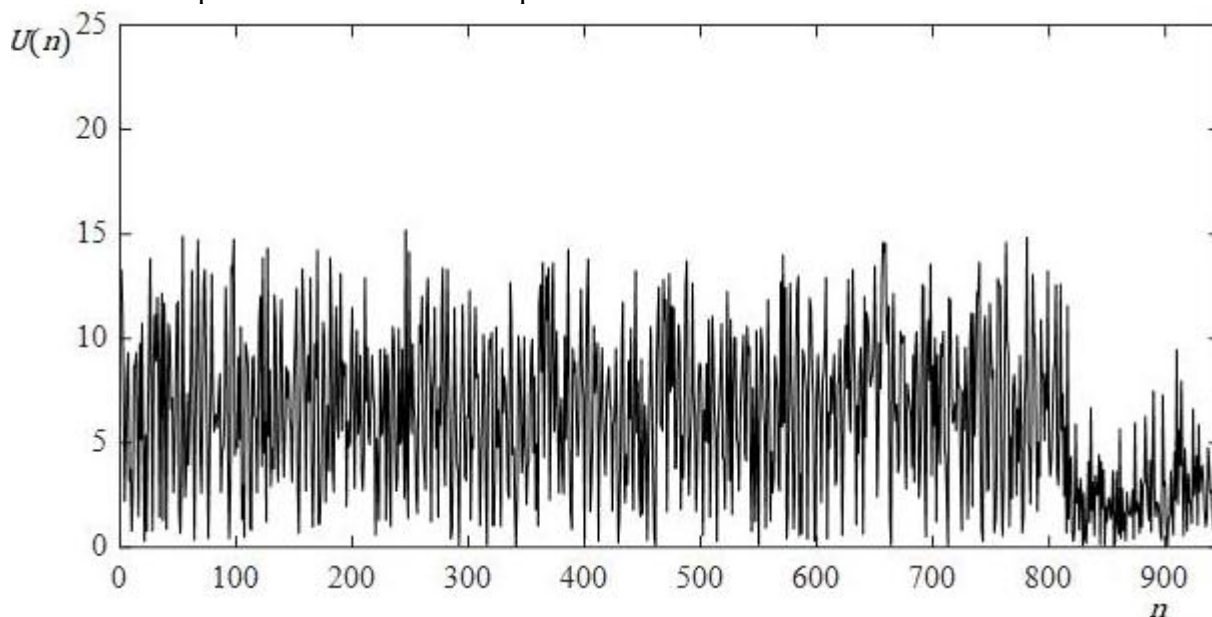


Рис. 3. Амплитуда процесса результата работы алгоритма подавления помех с постоянным порогом

Из рис. 3 видно, что произошло подавление всех шестнадцати импульсных помех до амплитуды сигнала. Также видно, что в результате работы алгоритма с постоянным порогом не происходило ограничения и подавления полезного сигнала, отраженного от цели.

На рис. 5 представлена амплитуда процесса $U(n)$ на входе алгоритма подавления помех и пороговый уровень со скользящим окном размером пять отсчетов

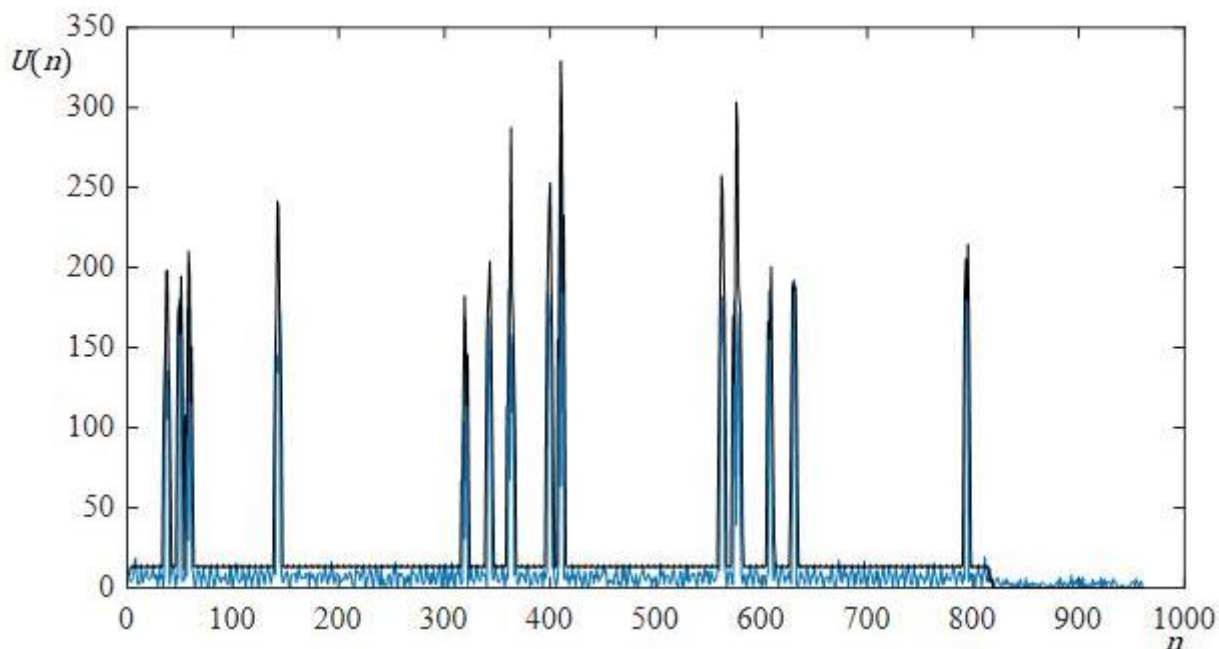


Рис. 5. Амплитуда процесса на входе алгоритма подавления помех и пороговый уровень со скользящим окном размером 5 отсчетов

На рис. 6 видно, что качество подавления помех при использовании алгоритма со скользящим окном длиной 5 отсчетов существенно ниже, чем при использовании постоянного порога. Это можно объяснить тем, что пороговый уровень, вычисляемый в относительно коротком скользящем окне завышается локально амплитудой помехи, а

уменьшить масштабирующий множитель порога нельзя из-за ограничения полезного сигнала. Ограничение полезного сигнала приводит к нелинейным искажениям и ухудшению измерений параметров движения объектов.

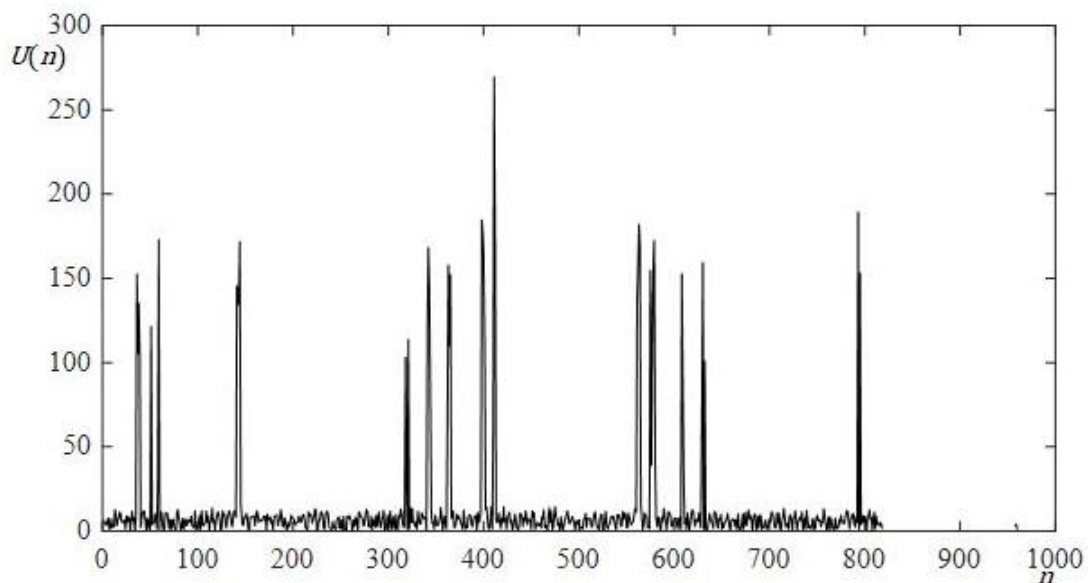


Рис. 6. Амплитуда процесса в результате работы алгоритма подавления помех со скользящим окном размером 5 отсчетов

На рис. 7 представлена амплитуда процесса $U(n)$ на входе алгоритма подавления помех с порогом, вычисленным в скользящем окне длиной 13 отсчетов

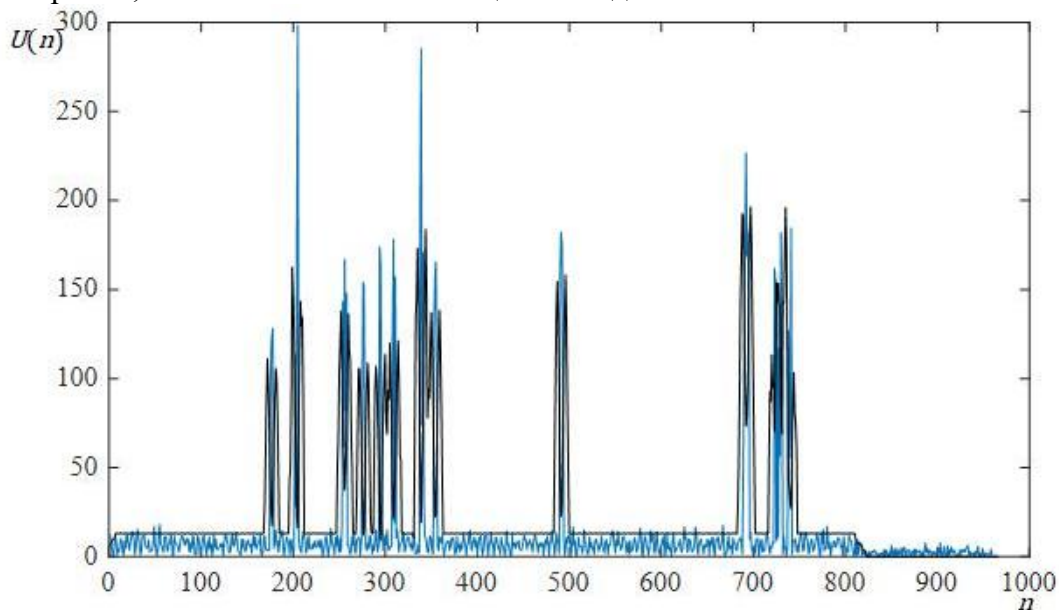


Рис. 7. Амплитуда процесса на входе алгоритма подавления помех с порогом, вычисленным в скользящем окне длиной 13 отсчетов

На рис.8 качество подавления помех лучше, чем при использовании алгоритма со скользящим окном длиной 5 отсчетов, но хуже, чем при использовании постоянного порога. Пороговый уровень, вычисляемый в скользящем окне длиной 13 отсчетов также завышается амплитудой помехи, что приводит к пропуску обнаружения отсчетов помехи и отсутствию ее подавления. Особенно часто такая ситуация возникает, когда несколько помех расположены близко друг к другу во времени. Тогда, процесс, соответствующий действию помехи, становится протяженным во времени и не может быть корректно обработан алгоритмом обнаружения и подавления со скользящим окном.

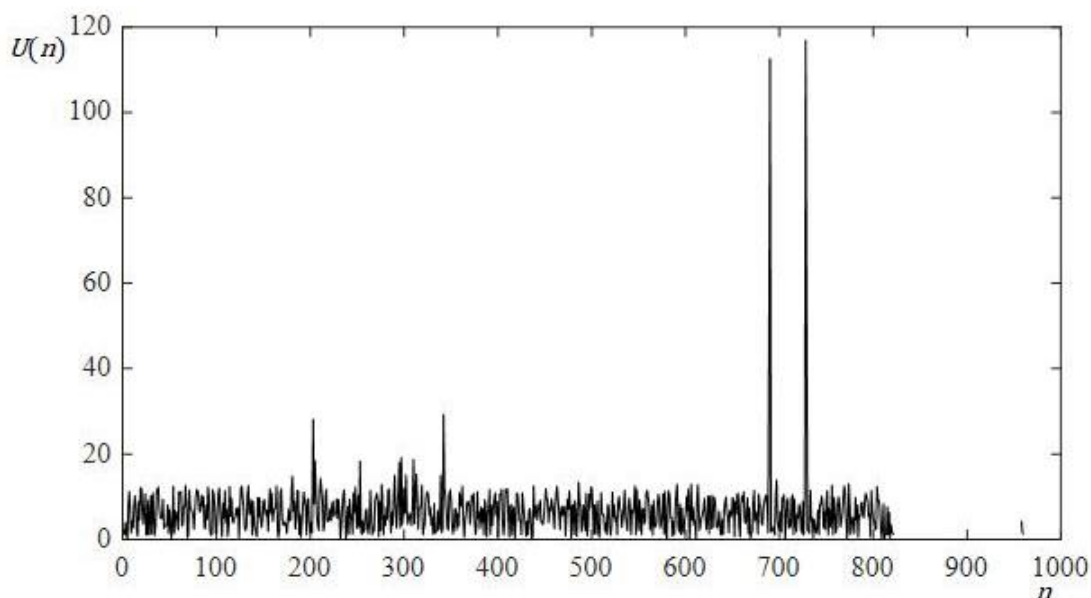


Рис. 8. Амплитуда процесса в результате работы алгоритма подавления помех со скользящим окном размером 13 отсчетов

Выводы

Методом математического моделирования были исследованы два алгоритма подавления импульсных помех в автомобильных радарх с непрерывным частотно-модулированным зондирующим сигналом. Алгоритм с постоянным порогом и с порогом, вычисленным в скользящем окне с весовой обработкой.

Результаты выполненного моделирования показали, что применение алгоритма с постоянным порогом является самым эффективным среди исследованных. Эффективность такого алгоритма может быть объяснена отсутствием зависимости порогового уровня от локальных значений амплитуд помех и их суммарной длительности во времени. Кроме того, при использовании постоянного порогового уровня удастся избежать нежелательного ограничения полезного сигнала.

Библиографический список

1. Masahiro Umehira, Yuya Makino, Tokeo Okuda, Xiaoyan Wang, Shigeki Takeda, Hiroshi Kuroda, "Inter-Radar Interference Analysis and Concept of Scalable Fast Chirp FMCW Radar for Automotive Applications", Conference: 2019 20th International Radar Symposium (IRS).

A.A. Kuzin, K.D. Pirogova

ALGORITHMS FOR SUPPRESSING ASYNCHRONOUS IMPULSE INTERFERENCE IN AUTOMOTIVE RADAR

Nizhny Novgorod state technical university n.a. R.E. Alekseev

This article is devoted to research the effectiveness of digital signal processing algorithms for suppressing asynchronous impulse interference in automotive radar. Without suppression of such interference, the signal-to-noise ratio decreases, which degrades the accuracy of estimating signal characteristics. The accuracy of measuring range, speed and angle depends on these characteristics. This report deals with algorithms for dealing with asynchronous impulse noise. Two noise suppression algorithms were considered, the results of modelling in the MATLAB are presented.

Keywords: car radar, asynchronous impulse interference, sliding threshold.

УДК 690.192; 623.624.9

А.И. Казьмин, П.А. Федюнин, Д.А. Рябов

ЭЛЕКТРОДИНАМИЧЕСКАЯ МОДЕЛЬ ОЦЕНКИ ВЕЛИЧИНЫ ОТСЛОЕНИЯ РАДИОПОГЛОЩАЮЩЕГО ПОКРЫТИЯ С ИСПОЛЬЗОВАНИЕМ ПОВЕРХНОСТНЫХ ЭЛЕКТРОМАГНИТНЫХ ВОЛН СВЧ ДИАПАЗОНА

Военный учебно-научный центр Военно-воздушных сил «Военно-воздушная академия имени профессора Н.Е. Жуковского и Ю.А. Гагарина» (г. Воронеж)

Представлена электродинамическая модель оценки величины отслоения радиопоглощающего покрытия с использованием поверхностных электромагнитных волн СВЧ диапазона. В отличие от существующих подходов в разработанной электродинамической модели измерений учитываются действительные и мнимые части комплексных диэлектрической и магнитной проницаемостей покрытия с учетом их частотной дисперсии, что повышает точность и достоверность оценки величины отслоения.

Ключевые слова: электродинамическая модель, отслоение, диэлектрическая и магнитная проницаемости, измерение, поверхностная электромагнитная волна.

Развитие методов и средств радиоволнового неразрушающего контроля радиопоглощающих материалов (РПМ) и покрытий (РПП) имеет важнейшее значение для повышения эффективности современных авиационных комплексов в целом [1,2].

Рассмотрим особенности применения метода поверхностных электромагнитных волн (ПЭМВ) для случая оценки величины отслоения однослойных магнитодиэлектрических покрытий (МДП) с диэлектрическими и магнитными потерями, с частотной дисперсией электрофизических параметров (ЭФП), от металлического основания.

На рис. 1 приведена геометрия однослойного МДП на металлической подложке. При этом учитывается, что в данной структуре может присутствовать протяженный дефект в виде воздушного отслоения толщиной d .

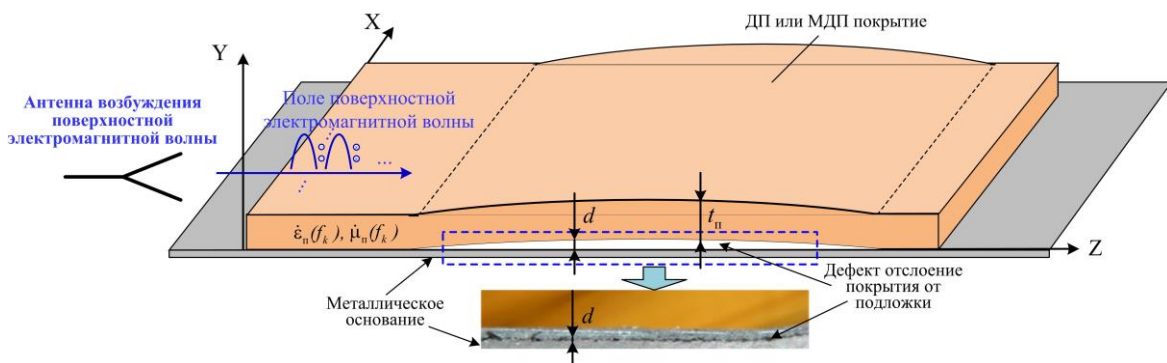


Рис. 1. Геометрия однослойного МДП при отслоении его от металлической подложки

Появление в МДП на металлической подложке протяженного дефекта в виде отслоения (длина отслоения в продольном направлении $1/3\lambda$ и более) можно представить, как

появление в системе «дополнительного слоя» с отличными от основного покрытия свойствами и рассматривать его как двухслойное покрытие «отслоение-покрытие». Обоснованность подобного подхода подтверждена теоретически и экспериментально в [3,4].

Вектор ЭФП покрытия представим пятикомпонентным, $\xi_n = \{\varepsilon'_n(f_k), \varepsilon''_n(f_k), \mu'_n(f_k), \mu''_n(f_k), t_n\}$, где $\varepsilon'_n(f_k), \mu'_n(f_k)$ действительные и мнимые $\varepsilon''_n(f_k), \mu''_n(f_k)$ части диэлектрической и магнитной проницаемостей соответственно, t_n – толщина покрытия, а вектор дефектов $\xi_d = \{d_s, d_{s+1}, \dots, d_M\}$ представим в виде одной величины отслоения d . Исходя из этого, комплексный коэффициент ослабления ПЭМВ будет иметь следующий вид $\hat{\alpha}(f, \xi_n, d)$.

Выражение для нахождения величины отслоения задается в виде целевой функции вида:

$$\hat{d} = \arg \min_{d \in d_{\text{дон}}} \rho(d, \alpha''(f, \xi_n, d)) = \sum_{k=1}^L |\hat{\alpha}_s(f_k, \xi_n, d) - \hat{\alpha}_t(f_k, \xi_n, d)|^2, \quad (1)$$

где $\hat{\alpha}_s(f, \xi_n, d)$ – экспериментальное значение коэффициента ослабления; $\hat{\alpha}_t(f, \xi_n, d)$ – теоретическое значение коэффициента ослабления.

Теоретические значения коэффициента ослабления $\hat{\alpha}(f, \xi_n, d)$ ПЭМВ рассматриваемого двухслойного покрытия «отслоение-покрытие» получали путем решения дисперсионного уравнения по известному вектору ЭФП ξ_n и пробным значениям величин отслоений d на частотах $f_k, k=1, 2, \dots, L$ [3,4]:

$$D[\hat{\alpha}(f, \xi_n, d); f_k] = 0. \quad (2)$$

Дисперсионное уравнение (2) составляли, используя метод «поперечного резонанса» [3,4]. Расчетная схема составления дисперсионного уравнения (2) приведена на рис. 2.

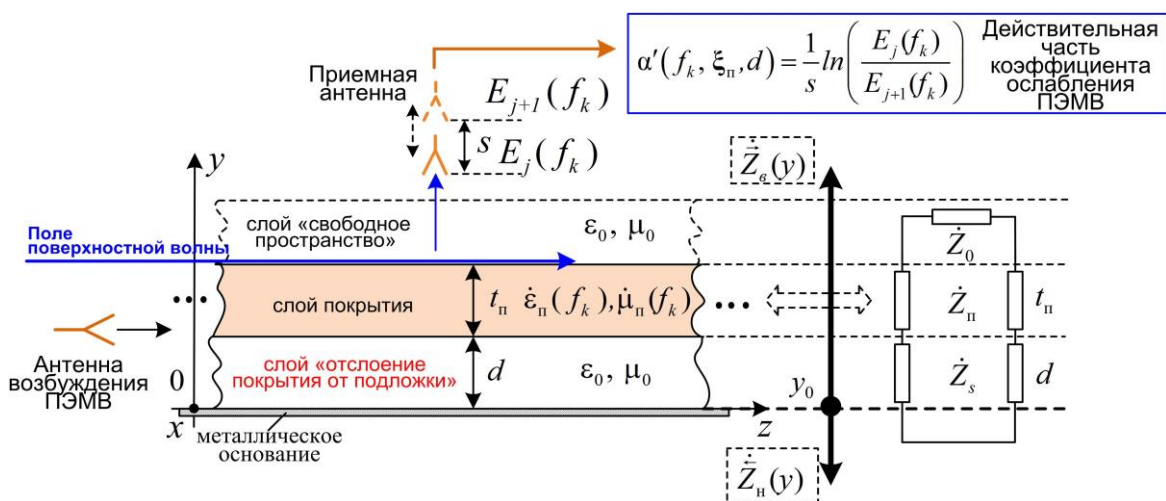


Рис. 2. Расчетная схема составления дисперсионного уравнения для ДП или МДП при отслоении от металлической подложки

В общем случае в данной структуре могут существовать ПЭМВ Е- и Н- типов. Дальнейшие теоретические и экспериментальные исследования приведены для ПЭМВ Е-типа. Между тем, физические особенности оценки отслоения с помощью ПЭМВ Н-типов аналогичны.

Уравнение «поперечного резонанса» для составления дисперсионного уравнения (2) имеет следующий вид [3,4]:

$$\dot{Z}_b(y) + \dot{Z}_n(y) = 0, \text{ при } \forall y, \quad (3)$$

где $\dot{Z}_b(y)$ и $\dot{Z}_n(y)$ – эквивалентные характеристические сопротивления «вверх» и «вниз» относительно произвольного опорного сечения y_0 (для удобства математических преобразований выбрано сечение y_0 между слоем «отслоение» и металлической подложкой).

Характеристическое сопротивление $\dot{Z}_n(y) = 0$, так как ниже опорного сечения y_0 – металлическая поверхность, а $\dot{Z}_b(y)$ определяется по рекуррентной формуле трансформации волновых сопротивлений [3,4]:

$$\dot{Z}_b \equiv \dot{Z}(s_{-n-0}) = \dot{Z}(s_{-n}) \frac{\dot{Z}_0 + j\dot{Z}(s_{-n}) \operatorname{tg}(\dot{q}t_n)}{\dot{Z}(s_{-n}) + j\dot{Z}_0 \operatorname{tg}(\dot{q}t_n)}, \quad \dot{Z}(s_{-n}) = \dot{Z}_s \frac{\dot{Z}_n + j\dot{Z}_s \operatorname{tg}(\dot{v}d)}{\dot{Z}_s + j\dot{Z}_n \operatorname{tg}(\dot{v}d)}, \quad (4)$$

где $\dot{Z}(s_{-n-0})$ – эквивалентное характеристическое сопротивление слоя «отслоение», слоя покрытия и области над покрытием (слой «свободное пространство»); $\dot{Z}(s_{-n})$ – эквивалентное характеристическое сопротивление слоя «отслоение» и слоя покрытия; $\dot{Z}_0, \dot{Z}_n, \dot{Z}_s$ – характеристические сопротивления на границе раздела между слоем покрытия и свободным пространством, в слое покрытия и в слое «отслоение», соответственно.

Выражения для $\dot{Z}_0, \dot{Z}_n, \dot{Z}_s$ имеют следующий вид [3,4]:

$$\dot{Z}_0 = -j \frac{\dot{\alpha}(f_k, \xi_n, d)}{\omega \epsilon_0}; \quad \dot{Z}_n = \frac{\dot{q}}{\omega \epsilon_0 (\epsilon'_n(f) - j\epsilon''_n(f))}; \quad \dot{Z}_s = \frac{\dot{v}}{\omega \epsilon_0},$$

где $\dot{\alpha}(f_k, \xi_n, d)$ – комплексный коэффициент ослабления ПЭМВ, $\dot{\alpha}(f_k, \xi_n, d) = \sqrt{\dot{\gamma}^2 - k_0^2}$, $\dot{\gamma}$ – продольная комплексная постоянная распространения ПЭМВ; k_0 – волновое число свободного пространства, $k_0 = 2\pi f / c$; \dot{q} – комплексное поперечное волновое число ПЭМВ в слое покрытия, $\dot{q} = \sqrt{\dot{k}_n^2 - k_0^2 - \dot{\alpha}^2(f, \xi_n, d) + \beta^2}$, \dot{k}_n – комплексное волновое число в слое покрытия, $\dot{k}_n = 2\pi f \sqrt{(\epsilon'_n(f) - j\epsilon''_n(f))(\mu'_n(f) - j\mu''_n(f))} / c$, c – скорость ЭМВ в свободном пространстве, $\epsilon'_n(f) - j\epsilon''_n(f)$ – комплексная диэлектрическая проницаемость слоя, $\mu'_n(f) - j\mu''_n(f)$ – комплексная магнитная проницаемость слоя, $\beta = m\pi/p$ – волновое число характеризующее распределение поля ПЭМВ по ширине покрытия p ; $\dot{v} = -j \cdot \dot{\alpha}(f, \xi_n, d)$ – комплексное волновое число в слое «отслоение»; j – мнимая единица.

Итоговое дисперсионное уравнение для покрытия с отслоением представим в следующем виде: $D[\dot{\alpha}(f_k, \xi_n, d); f_k] = 0 \equiv \dot{Z}(s_{-n-0}) = 0. \quad (5)$

Таким образом, полученное дисперсионное уравнение (5) фактически является электродинамической моделью, обеспечивающей однозначную оценку комплексного коэффициента ослабления ПЭМВ $\dot{\alpha}(f, \xi_n, d)$ для заданной частоты, ЭФП РПП, с учетом их частотной дисперсии и величины отслоения d .

Библиографический список

1. Лагарьков А.Н. Фундаментальные и прикладные проблемы стелс-технологий / А.Н. Лагарьков, М.А. Погосян // Вестник РАН. – 2003. – Т. – 73. – № 9. – С. 779–787.
2. Федюнин П. А. Способы радиоволнового контроля параметров защитных покрытий авиационной техники / П.А. Федюнин, А.И. Казьмин. – М.: Физматлит, 2013. – 190 с.
3. Казьмин А.И. Контроль дефектов в многослойных диэлектрических материалах СВЧ-методом / А.И. Казьмин, П.А. Федюнин // Заводская лаборатория. Диагностика материалов, 2020. – Том 86. – № 2. – С. 37–43.
4. Казьмин А.И. Оценка степени отслоения диэлектрических и магнетодиэлектрических покрытий с использованием поверхностных электромагнитных волн СВЧ диапазона / А.И. Казьмин, П.А. Федюнин // Дефектоскопия, 2020. – № 9. – С. 50–63.

A.I. Kaz'min, P. A. Fedyunin, D.A. Ryabov

ELECTRODYNAMIC MODEL FOR ESTIMATING OF EXFOLIATION A RADIO-ABSORBING COATING WITH SURFACE MICROWAVES

*Military Educational and Scientific Center of the Air Force
«N.E. Zhukovsky and Y.A. Gagarin Air Force Academy» (Voronezh)*

The electrodynamic model for evaluating the magnitude of delamination of a radio-absorbing coating using surface microwaves is proposed. In contrast to the existing approaches, the developed electrodynamic measurement model takes into account the real and imaginary parts of the complex dielectric permittivity and magnetic permeability of the coating with allowance for their frequency dispersion; this improves the accuracy and reliability of estimating the extent of exfoliation.

Keywords: electrodynamic model, delamination, permittivity, permeability, measurement, surface electromagnetic wave

ИМИТАЦИОННАЯ МОДЕЛЬ ДЛЯ ОЦЕНКИ ТОЧНОСТИ РЕКОНСТРУКЦИИ ЭЛЕКТРОФИЗИЧЕСКИХ ПАРАМЕТРОВ МНОГОСЛОЙНЫХ ДИЭЛЕКТРИЧЕСКИХ ПОКРЫТИЙ МЕТОДОМ ПОВЕРХНОСТНЫХ ЭЛЕКТРОМАГНИТНЫХ ВОЛН

*Военный учебно-научный центр Военно-воздушных сил «Военно-воздушная академия
имени профессора Н.Е. Жуковского и Ю.А. Гагарина» (г. Воронеж)*

Представлена имитационная модель, позволяющая оценивать точность реконструкции электрофизических и геометрических параметров многослойных диэлектрических покрытий многочастотным методом поверхностных электромагнитных волн. Имитационная модель реализована в виде набора подпрограмм в системе Matlab. Приведены результаты имитационного моделирования по реконструкции диэлектрических проницаемостей и толщин, одно- и двухслойных диэлектрических покрытий при различных значениях среднего квадратического отклонения уровня шума в измерительных данных.

Ключевые слова: многослойное диэлектрическое покрытие, многочастотный метод поверхностных электромагнитных волн, имитационная модель, электрофизические параметры.

Уменьшение мощности рассеиваемых радиоволн применяется для снижения радиолокационной заметности образцов вооружения, военной и специальной техники, создания антенн и антенных решеток с необходимыми свойствами, обеспечения электромагнитной совместимости и решении других задач. Один из способов уменьшения мощности рассеиваемых радиоволн заключается в применении диэлектрических покрытий (ДП), реализованных в виде радиопоглощающих покрытий (РПП). Их эффективность определяется электрофизическими и геометрическими параметрами (ЭФГП). От точности измерений ЭФГП зависят возможности объектов, в которых применяются РПП. Таким образом, разработка новых методов измерения ЭФГП многослойных ДП и РПП, в частности, является актуальной научной задачей [1,2].

В [3, 4] показано, что повышение эффективности контроля ЭФГП многослойных диэлектрических материалов и покрытий возможно путем применения многочастотного метода поверхностных электромагнитных волн (ПЭМВ). В качестве информативного параметра при этом выступает комплексный коэффициент ослабления поля ПЭМВ $\alpha(f, \xi_p, \xi_d)$, где f – частота зондирующего сигнала, $\xi_p = \{\dot{\epsilon}_i, \dots, \dot{\epsilon}_n, \dot{\mu}_{n+1}, \dots, \dot{\mu}_{2n}, b_{2n+1}, \dots, b_{3n}\}$, $i = 1, \dots, 3n$ – $3n$ -мерный вектор описывающий ЭФГП n -слойного покрытия, $\dot{\epsilon}_i, \dots, \dot{\epsilon}_n, \dot{\mu}_{n+1}, \dots, \dot{\mu}_{2n}$ – комплексные относительные диэлектрические и магнитные проницаемости, $b_{2n+1} = (t_{2n+1}/\lambda)_{2n+1}, \dots, b_{3n} = (t_{3n}/\lambda)_{3n}$ – относительные толщины слоёв n -слойного материала; t_{2n+1}, \dots, t_{3n} – толщины слоёв материала.

Цель работы – численное исследование потенциальных возможностей многочастотного радиоволнового метода поверхностных электромагнитных волн по точности реконструкции ЭФГП многослойных ДП в зависимости от значений их ЭФГП, уровня шума в измерительных данных и ширины полосы частот измерений.

При имитационном моделировании на ЭВМ выполняется совокупность расчётов, позволяющих имитировать процесс реконструкции относительных диэлектрических проницаемостей и толщин слоёв многослойного ДП по экспериментально полученным значениям коэффициентов ослабления поля ПЭМВ $\alpha_s(f, \xi)$.

Для моделирования разработана имитационная модель измерительного комплекса [3, 4], структурная схема которой приведена на рис. 1.

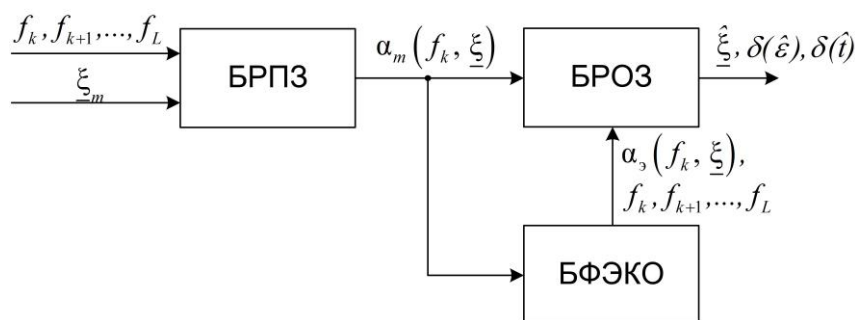


Рис. 1. Структурная схема имитационной модели

В имитационную модель входят три блока: блок решения прямой задачи БРПЗ, позволяющий определять теоретические значения $\alpha_T(f, \xi)$ коэффициентов ослабления поля для частот f_k , $k=1, 2, \dots, L$ и вектора ЭФГП $\xi_n = \{\varepsilon_i, \dots, \varepsilon_n, b_{n+1}, \dots, b_{2n}\}$ блок формирования экспериментальных значений $\alpha_3(f, \xi)$, $k=1, 2, \dots, L$, коэффициентов ослабления поля БФЭКО, имитирующий абсолютные погрешности $\Delta\alpha(f, \xi)$ измерений коэффициента ослабления поля; блок решения обратной задачи BROZ, реализующий нахождение минимума целевой функции.

В блоке БРПЗ рассчитываются теоретические значения коэффициентов ослабления $\alpha_T(f, \xi)$ многослойного диэлектрического ДП путём решения дисперсионного уравнения по известному вектору ЭФГП $\xi = \{\varepsilon_i, \dots, \varepsilon_n, b_{n+1}, \dots, b_{2n}\}$ на частотах f_k , $k=1, 2, \dots, L$ [3,4]:

$$D[\alpha_T(f_k, \xi); f_k] = 0. \quad (1)$$

В блоке BROZ реализуется представленный в [3] метод определения вектора ЭФГП $\xi = \{\varepsilon_i, \dots, \varepsilon_n, b_{n+1}, \dots, b_{2n}\}$, $i=1, \dots, 2n$, многослойных ДП. Метод основан на минимизации целевой функции, построенной по невязке между полученными экспериментально $\alpha_3(f, \xi)$ и вычисленными теоретически значениями $\alpha_T(f, \xi)$ коэффициентов ослабления поля на частотах f_k , $k=1, 2, \dots, L$:

$$\begin{aligned} \xi = \arg \min_{\xi \in \xi_{\text{доп}}} \rho(\xi) = \sum_{k=1}^L \left| \frac{\alpha_3(f_k, \xi) - \alpha_T(f_k, \xi)}{\alpha_T(f_k, \xi)} \right|^2 + \dots \\ \dots + \varpi \sum_i^{2n} q_i \left\| (\xi_i - \xi_{i0}) \right\|^2, \end{aligned} \quad (2)$$

где $\rho(\xi)$ – расстояние между значениями $\alpha_3(f, \xi)$ и $\alpha_T(f, \xi)$ в области допустимых значений $\xi \in \xi_{\text{доп}}$; q_i – вес, определяющий степень априорной информации о каждой составляющей вектора ЭФГП; ϖ – параметр регуляризации; ξ_i – i -я составляющая вектора ЭФГП; ξ_{i0} – найденное с учётом априорной информации о составляющей среднее значение i -й составляющей вектора ЭФГП из области допустимых значений $\xi_{i0} \in [\xi_{i\min}, \dots, \xi_{i\max}]$.

Ввиду множества случайных факторов, влияющих на точность реконструкции ЭФГП многослойных ДП, погрешность их измерения в блоке БФЭКО моделировали аддитивным некоррелированным гауссовым шумом $n(f_k)$ с нулевым математическим ожиданием $m(n(f_k)) = 0$ и заданным уровнем дисперсии σ_n^2 . Формировали искажённые значения результатов измерений частотной зависимости коэффициента ослабления поля:

$$\left. \begin{aligned} \alpha_o(f_k, \xi) &= \alpha_T(f_k, \xi) + n(f_k); \\ m(n(f_k)) &= 0; \\ k &= 1, 2, \dots, L, \end{aligned} \right\} \quad (3)$$

где $n(f_k)$ – шум, обусловленный как погрешностями самого физического прибора и другими измерительными погрешностями при проведении эксперимента, так и достоверностью выбранной электродинамической модели реальной физической ситуации; L – число частот, на которых проводятся измерения.

Таким образом, в соответствии с принятой моделью нормального распределения погрешностей измерений (3), имитировали погрешности измерений $\Delta\alpha(f_k, \xi)$.

Блоки структурной схемы имитационного моделирования (рис. 1) реализованы в виде подпрограмм в системе Matlab. Подпрограмма определения теоретических коэффициентов ослабления поля ПЭМВ для функционала (2) основана на решении дисперсионного уравнения (1) и реализована в виде функции на основе решателя fsolve.

Проведенные исследования показали, что при решении обратной задачи (2), помимо эффективных алгоритмов моделирования прямых задач для широкого диапазона частот, необходимо использовать соответствующие алгоритмы оптимизации целевой функции (функционала), которая в этом случае является невыпуклой и многопараметрической.

При числе слоев ДП больше трех прямые алгоритмы минимизации целевой функции (2) перестают давать устойчивые оценки ЭФГП и целесообразно использовать алгоритмы стохастической минимизации. Исходя из этого, подпрограммы определения ЭФГП ДП, основаны на применении генетического алгоритма и реализованы в виде функции «ga» приложения Global Search системы Matlab.

Имитационное моделирование включает следующее. По принятым точным значениям относительных диэлектрических проницаемостей $\hat{\epsilon}_1, \dots, \hat{\epsilon}_n$ и толщин t_1, \dots, t_n , $k = 1, 2, \dots, L$ многослойного ДП вычисляли теоретические значения коэффициентов ослабления $\alpha_T(f_k, \xi)$ для каждой из f_k фиксированных частот, решая дисперсионное уравнение (1).

В блоке БФЭКО формировали экспериментальные значения коэффициентов ослабления $\alpha_o(f_k, \xi)$, которые подставляли в качестве регистрируемой величины в целевую функцию (2). Далее в блоке БРОЗ оптимизировали выражение (2). Точность реконструкции относительных диэлектрических проницаемостей и толщин исследуемых диэлектрических покрытий исследовали при возрастающих значениях среднего квадратического отклонения (СКО) шума 0–0,008.

В качестве образцов для исследования выбраны ДП, схожие с реальными образцами приведенными в [5]. Также при имитационном моделировании изучали диэлектрики с относительной диэлектрической проницаемостью $\epsilon = 1, 2, \dots, 20$.

На рисунке 2а представлены зависимости средней относительной погрешности оценки диэлектрической проницаемости $\delta(\hat{\epsilon})$ однослойного ДП с толщиной $b = 1$ мм при увеличении значений диэлектрической проницаемости слоя от 1,2 до 20, для трех

значений СКО уровня шума: 0,001; 0,003; 0,005. На рисунке 2б представлено, как при этом меняется погрешность оценки толщины слоя $\delta(\hat{t})$. Коэффициент ослабления поля ПЭМВ определяли в диапазоне частот 9–10 ГГц с шагом по частоте 0,1 ГГц (количество частот $N=10$).

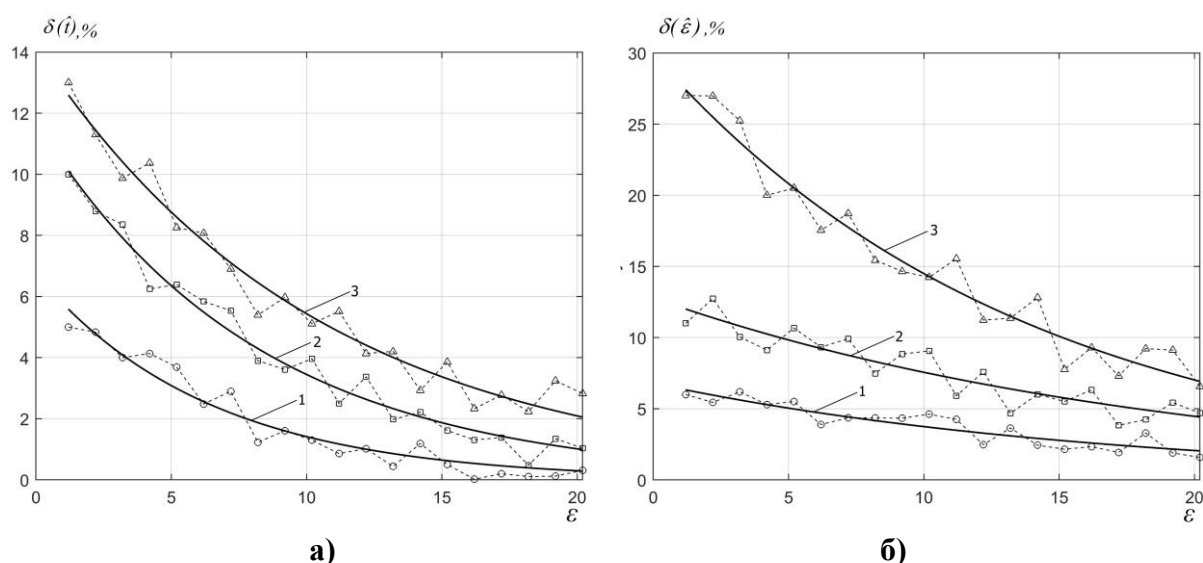


Рис. 2. Зависимости средней относительной погрешности оценки ЭФГП однослойного РПП при фиксированных значениях СКО уровня шума (1 – $\sigma=0,001$; 2 – $\sigma=0,003$; 3 – $\sigma=0,005$)

Анализ зависимостей рисунка 2а для однослойного ДП показывает, что относительная погрешность $\delta(\hat{\epsilon})$ оценки диэлектрической проницаемости слоя, при фиксированном значении уровня СКО шума и толщине слоя, уменьшается с возрастанием их величины. Относительная погрешность оценки толщины слоя $\delta(\hat{t})$, также уменьшается. При этом, точность оценки толщины слоя в 2–2,5 раза выше точности оценки диэлектрической проницаемости.

По результатам имитационного моделирования, метод дает оценки диэлектрической проницаемости и толщины для однослойного покрытия при СКО уровня шума $\sigma=0,003$, во всем диапазоне исследуемых значений диэлектрической проницаемости (1,2–20) с погрешностью не больше 10% с доверительной вероятностью 0,95.

Дисперсионное уравнение для двухслойного ДП усложняется тем, что появляется еще одна граница раздела сред между первым и вторым слоем покрытия.

Для численного эксперимента выбран наиболее сложный случай, когда двухслойное ДП является высококонтрастным (по диэлектрической проницаемости). Исследовано двухслойное ДП на основе диэлектриков с относительными диэлектрическими проницаемостями $\epsilon_1=2,7$ и $\epsilon_2=11,2$ с толщинами слоев 2 мм. Коэффициент ослабления поля ПМЭМВ, как и в случае однослойного покрытия, измеряли в диапазоне частот 9–10 ГГц с шагом по частоте 0,1 ГГц (количество частот $N=10$).

Увеличение числа оцениваемых ЭФГП приводит к возрастанию их относительных погрешностей. Погрешности оценок параметров слоев двухслойного РПП не совпадают с оценками для однослойного покрытия, однако, сохраняется общая закономерность, установленная для базового слоя – погрешности оценок диэлектрических проницаемостей и толщин слоев меньше при больших значениях диэлектрической проницаемости и электрической толщины слоя.

Метод дает устойчивые оценки относительных погрешностей ЭФГП слоев двухслойного ДП не больше 10 %, при уровне СКО шума $\sigma=0,003$ и более, с доверительной вероятностью 0,95. При этом экспериментальные значения погрешностей

оценок диэлектрических проницаемостей $\delta(\hat{\varepsilon}_1)$ и $\delta(\hat{\varepsilon}_2)$ слоев и их толщин $\delta(\hat{t}_1)$ и $\delta(\hat{t}_2)$ для покрытия « $\varepsilon_1 - \varepsilon_2$ » составляют 3% и 7 %, 2% и 6%, соответственно, а для покрытия « $\varepsilon_2 - \varepsilon_1$ » 4% и 5 %, 8% и 0,1 %, соответственно.

Отличие ЭФГП оценок отдельных слоев от ЭФГП однослойного покрытия, объясняется появлением дополнительных корреляционных связей между ними.

Таким образом, разработана имитационная модель, позволяющая оценивать точность реконструкции структуры электрофизических и геометрических параметров многослойных ДП в зависимости от их электрофизических и геометрических параметров, учитывающая значение СКО уровня шума в измерительных данных и ширину полосы частот измерений.

Библиографический список

1. Лагарьков А.Н. Фундаментальные и прикладные проблемы стелс-технологий / А.Н. Лагарьков, М.А. Погосян // Вестник РАН. – 2003. – Т. – 73. – № 9. – С. 779–787.
2. Федюнин П. А. Способы радиоволнового контроля параметров защитных покрытий авиационной техники / П.А. Федюнин, А.И. Казьмин. – М.: Физматлит, 2013. – 190 с.
3. Казьмин А.И. Восстановление структуры электрофизических параметров многослойных диэлектрических материалов и покрытий по частотной зависимости коэффициента ослабления поля поверхностной электромагнитной волны / А.И. Казьмин, П.А. Федюнин // Измерительная техника. – 2019. – № 9. – С. 39–45.
4. Казьмин А.И. Оценка точности реконструкции электрофизических и геометрических параметров многослойных диэлектрических покрытий многочастотным радиоволновым методом поверхностных медленных электромагнитных волн / А.И. Казьмин, П.А. Федюнин // Измерительная техника. – 2020. – 8. С. – 51–58.
5. Казанцева Н.Е. Перспективные материалы для поглотителей электромагнитных волн сверхвысококачастотного диапазона / Н.Е. Казанцева, Н.Г. Рывкина, И.А. Чмутин // Радиотехника и электроника. – 2003. – том 48. – № 2. – С. 196–209.

A.I. Kaz'min, P. A. Fedyunin, D.A. Ryabov,

THE SIMULATION MODEL FOR ESTIMATING THE ACCURACY OF RECONSTRUCTION OF ELECTROPHYSICAL PARAMETERS OF MULTILAYER DIELECTRIC COATINGS BY THE METHOD OF SURFACE ELECTROMAGNETIC WAVES

*Military Educational and Scientific Center of the Air Force
«N.E. Zhukovsky and Y.A. Gagarin Air Force Academy» (Voronezh)*

The simulation model is presented that makes it possible to evaluate of the accuracy of reconstruction of the structure of the electrophysical and geometric parameters of multilayer dielectric coatings depending on their electrophysical and geometric parameters, taking into account the value of the noise level in the measurement data and the measurement frequency bandwidth. The simulation model is implemented as a set of subroutines in the Matlab system. The results of simulation investigations of reconstruction of the structure of relative permittivities and thicknesses of single-layer and double-layer dielectric coatings with different thicknesses, with different values of the standard deviation (RMS) of the noise level in the measured attenuation coefficients of the surface slow electromagnetic wave are presented.

Keywords: multilayer dielectric coating, multifrequency method of surface electromagnetic waves, simulation model, electrophysical parameters.

АЛГОРИТМ ОБРАБОТКИ РЕЗУЛЬТАТОВ ИЗМЕРЕНИЯ КОЭФФИЦИЕНТОВ. ОСЛАБЛЕНИЯ ПОЛЯ ПОВЕРХНОСТНОЙ ЭЛЕКТРОМАГНИТНОЙ ВОЛНЫ ПРИ КОНТРОЛЕ ДЕФЕКТОВ В РАДИОПОГЛОЩАЮЩИХ ПОКРЫТИЯХ

Военный учебно-научный центр Военно-воздушных сил «Военно-воздушная академия имени профессора Н.Е. Жуковского и Ю.А. Гагарина» (г. Воронеж)

Разработан алгоритм оценки величины отслоения покрытия на основе поиска максимальной совместной плотности вероятности коэффициента ослабления поля поверхностной электромагнитной волны, измеренного на наборе частот. Обоснован и введен статистический предел разрешения двух величин отслоений, что позволяет оценить возможность различения двух близких величин отслоений в зависимости от ширины полосы частот измерений коэффициента ослабления, их количества и отношения «сигнал–шум».

Ключевые слова: отслоение, покрытие, поверхностная электромагнитная волна, коэффициент ослабления поля, диэлектрическая и магнитная проницаемости, частотная дисперсия, максимум совместной плотности распределения вероятностей, статистический предел разрешения.

Магнитодиэлектрические покрытия (МДП), работающие в диапазоне сверхвысоких частот (СВЧ), широко применяются в различных областях науки и техники. В отдельное важнейшее самостоятельное направление выделяется применение МДП, реализованных в виде радиопоглощающих покрытий (РПП), для снижения радиолокационной заметности современных образцов авиационных комплексов (АК) [1,2].

Эффективность применения РПП определяется их надежной адгезией к обшивке образца АК. При недостаточной адгезии в ходе эксплуатации высокие внутренние напряжения приводят к отслоениям РПП, что приводит к увеличению эффективной поверхности рассеяния образца АК и снижению его боевых возможностей [1,2].

Представленный в [3] метод оценки толщин протяженных дефектов типа отслоение многослойного диэлектрического покрытия от подложки и расслоений между слоями, основан на определении вектора оцениваемых геометрических параметров дефектов (их толщин) $\xi_d = \{d_s, d_{s+1}, \dots, d_M\}$ многослойного материала путем минимизации целевой функции, построенной по невязке между экспериментально полученными $\alpha'_s(f_k, \xi_p, \xi_d)$ и вычисленными теоретическими значениями действительной части коэффициентов ослабления $\alpha'_t(f_k, \xi_p, \xi_d)$ поля ПЭМВ на частотах $f_k, k=1, 2, \dots, L$:

$$\hat{\xi}_d = \arg \min_{\xi \in \xi_{\text{доп}}} \rho(\xi_d) = \sum_{k=1}^L |\alpha'_s(f_k, \xi_p, \xi_d) - \alpha'_t(f_k, \xi_p, \xi_d)|^2, \quad (1)$$

где $\rho(\xi_d)$ – расстояние между экспериментально полученными $\alpha'_s(f_k, \xi_p, \xi_d)$ и вычисленными теоретическими значениями $\alpha'_t(f, \xi_p, \xi_d)$ коэффициентов ослабления поля ПЭМВ в области допустимых значений $\xi \in \xi_{\text{доп}}$; f – частота зондирующего сигнала $\xi_p = \{\dot{\epsilon}_i, \dots, \dot{\epsilon}_n, \dot{\mu}_{n+1}, \dots, \dot{\mu}_{2n}, b_{2n+1}, \dots, b_{3n}\} \quad i=1, \dots, 3n$ – $3n$ -мерный вектор, описывающий ЭФП n -слойного покрытия, $\dot{\epsilon}_i, \dots, \dot{\epsilon}_n, \dot{\mu}_{n+1}, \dots, \dot{\mu}_{2n}$ – комплексные относительные диэлектрические и магнитные проницаемости, $b_{2n+1} = (t_{2n+1}/\lambda)_{2n+1}, \dots, b_{3n} = (t_{3n}/\lambda)_{3n}$ – относительные толщины слоёв n -слойного материала; t_{n+1}, \dots, t_{2n} – толщины слоёв материала; λ – длина волны

зондирующего сигнала, $\xi_d = \{d_s, d_{s+1}, \dots, d_M\}$ – вектор геометрических параметров дефектов (величин отслоений и расслоений в покрытии).

Предварительные исследования метода контроля отслоения РПП от металлической подложки d методом поверхностных электромагнитных волн (ПЭМВ), представленные в [3] и основанные на минимизации целевой функции (1), не дают ответ, насколько будет точным результат, с учетом неопределенностей в измерениях коэффициента ослабления поля ПЭМВ $\alpha'_o(f, \xi_n, d)$, d – величина отслоения РПП [3].

Исходя из этого, дальнейшее развитие метода заключается в разработке алгоритма обработки результатов измерений коэффициента ослабления ПЭМВ, обеспечивающего оценку величины отслоения РПП от подложки \hat{d} , с учетом неопределенностей в измерительных данных с наименьшей погрешностью.

Разработан алгоритм оценки величины отслоения РПП на основе поиска максимальной совместной плотности вероятности (МСПРВ) коэффициента ослабления поля ПЭМВ $\alpha'_o(f, \xi_n, d)$, измеренного на наборе частот [4].

Алгоритм включает следующую последовательность действий.

1. Определение диапазона возможных значений величин отслоений $d \in \{d_1, d_2, \dots, d_N\}$.
2. Оптимизация набора частот и ширины полосы частот измерений в соответствии со следующим правилом:

$$\langle F \rangle = \arg \max_f \{Y\}, \quad (2)$$

где

$$F = \{f_1, f_2, \dots, f_L\},$$

$$Y = \frac{1}{LN} \sum_{i=1}^L \sum_{n=1, r=1}^N |\alpha'_o(f_k, \xi_n, d_n) - \alpha'_o(f_k, \xi_n, d_r)|^2,$$

где L – количество частот измерений, N – количество возможных величин отслоений, d_n, d_r – близкие значения величин отслоений на частоте f_k .

3. Измерение вектора экспериментальных коэффициентов ослабления \mathbf{A}_o , где

$$\mathbf{A}_o = [\dot{\alpha}_o(f_1, \xi_n, d), \dot{\alpha}_o(f_2, \xi_n, d), \dots, \dot{\alpha}_o(f_L, \xi_n, d)]^T.$$

4. Минимизация функционала $Z(d_n)$ с учетом всех возможных величин отслоений:

$$Z(d_n) = \frac{1}{L} |\mathbf{A}_o - \mathbf{A}_t|^2, \quad (3)$$

где $\mathbf{A}_t = [\dot{\alpha}_t(f_1, \xi_n, d), \dot{\alpha}_t(f_2, \xi_n, d), \dots, \dot{\alpha}_t(f_L, \xi_n, d)]^T$ – вектор теоретических коэффициентов ослабления.

За истинную величину отслоения \hat{d} принимают ту, которая обеспечивает минимум функционала (3).

Разработанный алгоритм позволяет оценить статистический предел разрешения отслоения покрытия, т.е. выявить минимальную разность двух различных величин отслоений $\varpi = d_1 - d_2$. Введение данного понятия позволяет оценить метод в возможности различения двух близких величин отслоений d_1 и d_2 , что очень важно для практического применения. Оценка статистического предела разрешения обеспечивает необходимую

исходную информацию для выбора ширины полосы частот измерений и их количества для оценки отслоения.

Неопределенности в измерении коэффициентов ослабления A_k , включают шум в измерительной системе, погрешности в определении номинальных значений ЭФП слоев покрытия, калибровка измерительной системы и др. Исходя из этого, экспериментальные значения действительной части коэффициента ослабления ПЭМВ $\alpha'_3(f, \xi_n, d)$, как функции частоты и величины отслоения d , наблюдаются в эксперименте при наличии аддитивных случайных помех, и были представлены в виде [4]:

$$\alpha'_3(f_k, \xi_n, d) = \alpha'_t(f_k, \xi_n, d) + n(f_k); \quad k = 1, 2, \dots, L, \quad (4)$$

где $\alpha'_3(f, \xi_n, d)$ – экспериментальное значение действительной части коэффициента ослабления поля ПМЭМВ, $\alpha'_t(f, \xi_n, d)$ – истинное (теоретическое) значение коэффициента ослабления поля ПМЭМВ; $n(f_k)$ – функция неопределенности измерения, которая представляет собой шум, воздействующий на k -измерение.

Так как шум имеет нормальное распределение с нулевым математическим ожиданием, среднее значение оценки величины отслоения и предел его статистического разрешения, а также их среднего квадратического отклонения (СКО), являются функциями уровня шума. Среднее значение отслоения \hat{d} будет группироваться вокруг истинного значения отслоения d в интервале, определяемым СКО. При этом, вероятность того, что среднее значение оценки отслоения \hat{d} действительно находится в окрестностях своего истинного значения d , напрямую связана с тем, как определяется эта окрестность.

Предел статистического разрешения, который вводится в данной статье, основан на определении этой окрестности в зависимости от среднего квадрата ошибки определения величины отслоения $\Delta_d = \langle \hat{d} - d \rangle^2$, которую обеспечивает функционал оценки МСПРВ (3) и доверительной вероятности.

Среднее значение Δ_d по всем возможным расслоениям $d \in \{d_1, d_2, \dots, d_N\}$ представим следующим образом:

$$\Delta_d = \frac{1}{N} \sum_{n=1}^N \langle \hat{d}_n - d_n \rangle^2, \quad (5)$$

где \hat{d}_n – n -величина отслоения, полученная при минимизации функционала МСПРВ (3).

Предел статистического разрешения ϖ для полученной оценки Δ_d с учетом заданной доверительной вероятности p , $0 \leq p \leq 1$ представим в следующем виде:

$$\varpi = 2\zeta \sqrt{\Delta_d}, \quad (6)$$

при этом ζ определяется путем решения уравнения

$$p = \frac{1}{\sqrt{2\pi}} \int_{x-\zeta}^{x+\zeta} e^{-\frac{x^2}{2}} dx, \quad (7)$$

где p – вероятность того, что средняя величина отслоения \hat{d} лежит в пределах разрешения, около истинной величины отслоения d , x – стандартизированное значение Δ_d .

Предел статистического разрешения в (5) определяется исходя из предположения, что средний квадрат ошибки $1/N \sum_{n=1}^N |\hat{d}_n - d_n|^2$, подчиняется нормальному закону распределения с нулевым математическим ожиданием и с СКО, равным $\sqrt{\Delta_d}$. Это предположение оправдано, поскольку шум подчиняется нормальному закону распределения.

Таким образом, фактически, предел статистического разрешения, определяемый по выражению (6), является доверительным интервалом оценки величины отсложения, полученной при измерении. Исходя из этого, уменьшая Δ_d или доверительную вероятность, можно получить более высокие величины разрешения. Естественно, на практике, требуются доверительные вероятности 0,95 и более.

На рисунке 1 представлены зависимости среднего квадрата ошибки Δ_d оценки величины отсложения РПП при возрастании значений отношения «сигнал-шум» (ОСШ), при возрастании ширины полосы частот измерений и возрастании количества частот измерений. В качестве образца РПП со следующими характеристиками [5]: $\varepsilon' = 20,45$, а $\varepsilon'' = 0,73$, дисперсия комплексной магнитной проницаемости РПП описывается зависимостями:

$$\varepsilon'(f_k) = 1,56 \exp(-3.867 \cdot 10^{-11} f_k); \quad \varepsilon''(f_k) = 1,299 \exp(-4.659 \cdot 10^{-11} f_k).$$

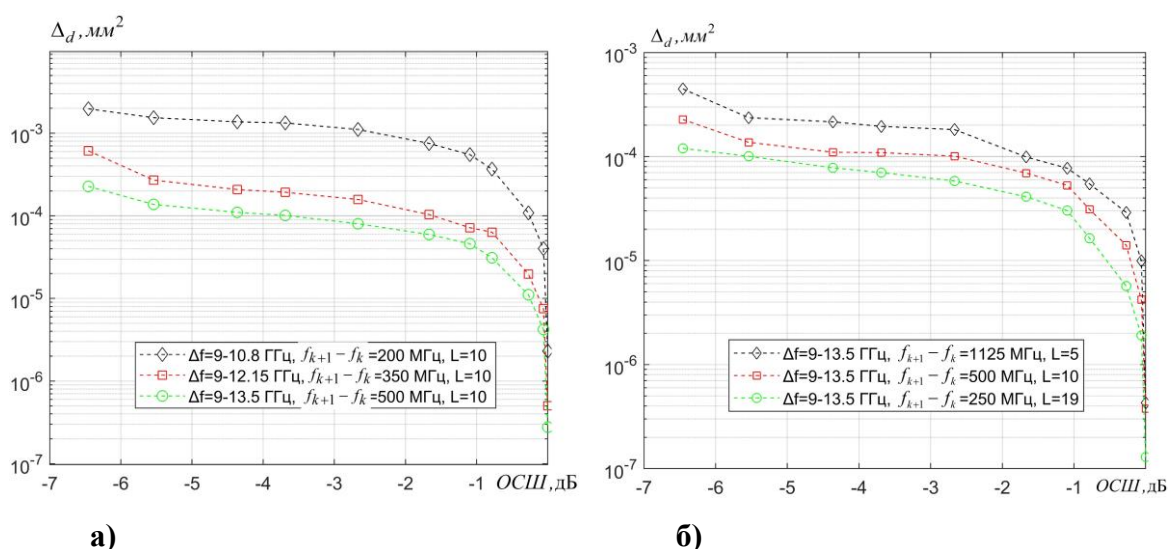


Рис. 1. Зависимости среднего квадрата ошибки Δ_d оценки отсложения в РПП при возрастании значений отношения «сигнал-шум» при возрастании ширины полосы частот измерений (а) и при возрастании количества частот измерений (б)

Анализ зависимостей рисунка 1 показывает очевидный результат, что средний квадрат ошибки Δ_d оценки величины отсложения уменьшается при возрастании ОСШ, увеличения ширины Δf полосы частот измерений и их количества. При этом уменьшение Δ_d за счет увеличения ширины полосы частот измерений достигается быстрее, чем при увеличении количества частот. Проведенные численные эксперименты показали, что при ширине полосы частот $\Delta f = 9 - 13,5 \text{ ГГц}$, предельный прирост повышения точности обеспечивается при измерениях на 19 частотах.

На основании полученных зависимостей оценки Δ_d оценивали статистический предел разрешения ϖ , как функцию от значений доверительной вероятности. На рисунке

2 приведены зависимости предела разрешения для ОСШ $-0,27$ дБ и $-2,67$ дБ для диапазона $9\text{--}13,5$ ГГц при измерениях на $L=19$ частотах.

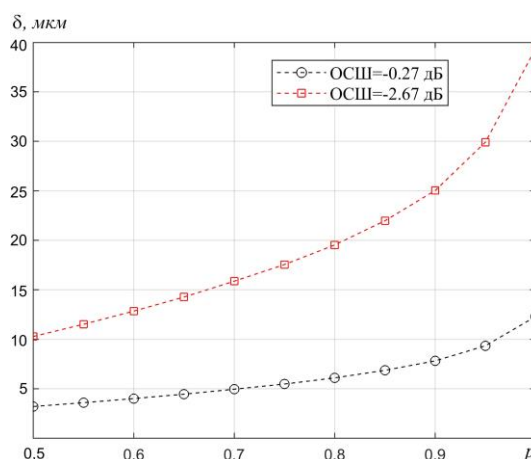


Рис. 2. Зависимость предела разрешения от доверительной вероятности

Анализ зависимостей рис. 2 показывает, что с увеличением ОСШ, разрешение увеличивается для всех значений доверительной вероятности. Это в основном связано с тем, что Δ_d монотонно уменьшается при возрастании ОСШ (рисунок 1). Анализ зависимостей рис. 2, показывает, что при доверительной вероятности 0,95 минимально достижимые уровни разрешения составляют около 9 и 30 мкм при ОСШ $-0,27$ дБ и $-2,67$ дБ, соответственно.

Оценка МСПРВ при ОСШ $-2,67$ дБ, с доверительной вероятностью 0,95 не будет чувствительна к изменению отслоения с шагом менее чем 30 мкм. Если требуется уменьшить разрешение менее 30 мкм, необходимо увеличить ОСШ в системе. Например, из анализа графиков рисунка 4, следует, что увеличение ОСШ в 13 раз до $-2,67$ дБ обеспечивает разрешение около 9 мкм с доверительной вероятностью 0,95.

Таким образом, разработан алгоритм оценки величины отслоения диэлектрических и магнетодиэлектрических покрытий \hat{d} на основе поиска максимума совместной плотности распределения вероятности, позволяющий оценить нижний предел среднего квадрата ошибки Δ_d оценки отслоения. Введено понятие о предельном разрешении по величине отслоения покрытия, что позволяет оценить метод в возможности различении двух близких величин отслоений.

Библиографический список

1. Лагарьков А.Н. Фундаментальные и прикладные проблемы стелс-технологий / А.Н. Лагарьков, М.А. Погосян // Вестник РАН. – 2003. – Т. – 73. – № 9. – С. 779–787.
2. Федюнин П. А. Способы радиоволнового контроля параметров защитных покрытий авиационной техники / П.А. Федюнин, А.И. Казьмин. – М.: Физматлит, 2013. – 190 с.
3. Казьмин А.И. Контроль дефектов в многослойных диэлектрических материалах СВЧ-методом / А.И. Казьмин, П.А. Федюнин // Заводская лаборатория. Диагностика материалов, 2020. – Том 86. – № 2. – С. 37–43.
4. Казьмин А.И. Оценка степени отслоения диэлектрических и магнетодиэлектрических покрытий с использованием поверхностных электромагнитных волн СВЧ диапазона / А.И. Казьмин, П.А. Федюнин // Дефектоскопия, 2020. – № 9. – С. 50–63.
5. Ufimtsev P.Ya. New Results for the Properties of TE Surface Waves in Absorbing Layers / P.Ya. Ufimtsev, R.T. Ling // IEEE Transactions on Antennas and Propagation. – vol. 49. – Issue 10. – pp. 1445–1452. – Oct 2001.

A.I. Kaz'min, P. A. Fedyunin, D.A. Ryabov, V.A. Manin

**MEASUREMENT RESULTS PROCESSING ALGORITHM OF FIELD ATTENUATION
COEFFICIENT OF A SURFACE ELECTROMAGNETIC WAVE UNDER THE
CONTROL OF DEFECTS IN RADIO-ABSORBING COATINGS**

*Military Educational and Scientific Center of the Air Force
«N.E. Zhukovsky and Y.A. Gagarin Air Force Academy» (Voronezh)*

An algorithm is developed for estimating the extent of coating exfoliation based on the search for the maximum joint probability density of the coefficient of attenuation of the field of a surface electromagnetic wave measured at a set of frequencies. We introduce and substantiate a statistical limit of the resolution of two values of exfoliation; this makes it possible to evaluate the possibility of discriminating between two close values of exfoliation depending on the bandwidth of attenuation coefficient measurements, their number, and the signal-to-noise ratio

Keywords: exfoliation, coating, surface electromagnetic wave, field attenuation coefficient, permittivity and permeability, frequency dispersion, maximum of joint probability distribution density, statistical resolution limit.

ИССЛЕДОВАНИЕ ВЛИЯНИЯ СТРУКТУРЫ УГЛЕКОМПОЗИТНЫХ МАТЕРИАЛОВ НА ПАРАМЕТРЫ РУПОРНЫХ АНТЕНН СВЧ ДИАПАЗОНА

Волжский государственный университет водного транспорта

В работе рассматривается влияние структуры углекомпозитного материала на параметры изготовленных из него высокочастотных антенных устройств. При помощи компьютерного моделирования исследовано влияние проводимости материала стенок рупорной антенны на ее коэффициент усиления. Экспериментально показано наличие анизотропных свойств проводимости композиционного материала в зависимости от способа его применения. Были созданы рупорные антенны L- и С- диапазонов частот из углекомпозитных материалов с различной структурой и графеносодержащим эпоксисвязующим веществом. Исследование поляризационных характеристик этих антенн показало, что наличие анизотропной проводимости стенок волноводной части антенн, приводит к изменению их радиочастотных параметров.

Ключевые слова: рупорная антенна, углекомпозит, графен, поляризация.

В настоящее время при проектировании и производстве современных высокотехнологичных устройств все чаще применяются материалы на основе углерода – углекомпозиты, поскольку они обладают высокой температурной стабильностью, повышенной прочностью, относительно низким удельным весом и широким диапазоном значений проводимости.

Ранее проведенные исследования показали возможность применения углекомпозитных материалов (УКМ) для создания радиочастотных компонентов антенно-фидерных устройств различных конструкций в широком частотном диапазоне [1].

Для изучения влияния анизотропии проводимости материалов на параметры антенн были исследованы рупорные антенны, изготовленные из углекомпозита различной структуры. В качестве образца-эталона была выбрана антенна L- и С- диапазонов (центральные частоты – 1,6 и 5 ГГц), состоящие из последовательно соединенных рупорной части и круглого волновода, изготовленных из латуни с серебряным покрытием

В точном соответствии с внутренними размерами металлических аналогов были изготовлены заготовки-матрицы, для последующей укладки на них УКМ.

Для изготовления модельных образцов была выбрана углекомпозитная нить марки Zoltek PX 35, обладающая электрическим сопротивлением $1.55 \cdot 10^{-5}$ Ом·м [2], и ткань, сплетенная из этой нити с ячейкой 20x20 мм.

Далее в программном пакете ANSYS HFSS была построена компьютерная модель рассматриваемой рупорной антенны С- диапазона и проведено исследование влияния проводимости материала, из которого она изготовлена, на ее коэффициент усиления. Расчеты выполнялись с использованием лицензионной версии программной системы на кафедре «Физика и техника оптической связи» образовательного научно-исследовательского института ядерной энергетики и технической физики им. академика Ф.М. Митенкова НГТУ им. Р.Е. Алексеева.

Полученная характеристика (нормированная) представлена на Рис.1 сплошной линией. Штриховой линией указано значение проводимости (ρ) для нити Pх35, штрихпунктирной линией – для алюминия, а точечной – для меди. Для антенны L-диапазона – результат моделирования качественно аналогичен.

При рассмотрении характеристик зависимости усиления антенны от значения проводимости материала, из которого она изготовлена, видно, что удельное сопротивление выбранного углекомпозита Zoltek PX 35 находится в рабочей зоне и

позволяет использовать его для создания антенно-фидерных устройств. Для увеличения прочности, придания проводящих свойств связующему и фиксации углекомпозитного материала на заготовках-матрицах использовалась эпоксидная смола с добавлением графеносодержащих структур. Нить и ткань наносились на заготовку методом круговой намотки с пропиткой графеносодержащим эпоксисвязующим. После отвердевания связующего вещества заготовка извлекалась. Таким способом были получены два модельных образца углекомпозитных рупорных антенн (из нити и ткани) диапазона 5 ГГц (рис. 2) и модельный образец УКМ антенны на 1,6 ГГц из нити.

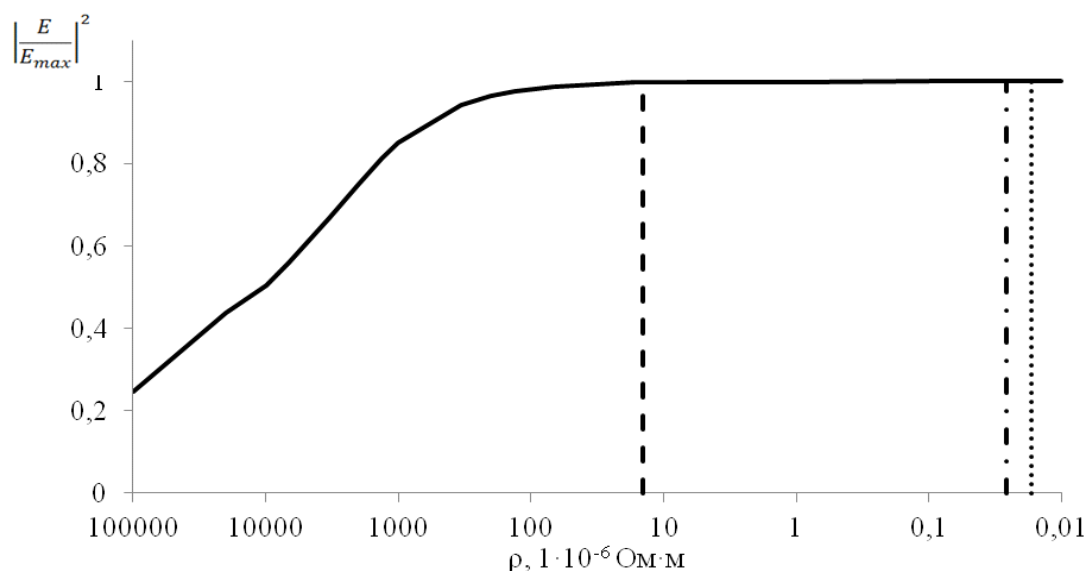


Рис. 1. Характеристика зависимости КУ рупорной антенны С- диапазона от удельного сопротивления материала, из которого она изготовлена.

Для получения эллиптической поляризации излучаемых антеннами волн использовалось поляризационное устройство, состоящее из 6 подвижных металлических штырей, установленное на волноводную часть антенн.

Для исследования поляризационных характеристик всех антенн в качестве излучателя использовалась эталонная рупорная антенна П6-23А, запитываемая от генератора R&S SMB100А. Исследуемые антенны устанавливались на штатном опорно-поворотном устройстве эталонной антенны с возможностью вращения антенн относительно своей оси. Величина принятого сигнала фиксировалась измерителем мощности Я2М-66 с термоэлектрической головкой.



Рис. 2. Рупорные антенны С- диапазона с возбуждающим устройством и поляризатором из УКМ (слева и в центре) и из металла (справа).

Измерения характеристик антенн на двух поляризациях волн (П1 и П2) проводились при помощи стандартных методик [3]. На рисунках 3 и 4 представлены обобщенные результаты, характеризующие зависимость коэффициентов эллиптичности

от частоты для всех антенн на двух поляризациях. Сплошной линией показана характеристика для металлической антенны, штриховой линией для углекомпонитной антенны из ткани, штрихпунктирной линией – углекомпонитной антенны из нити.

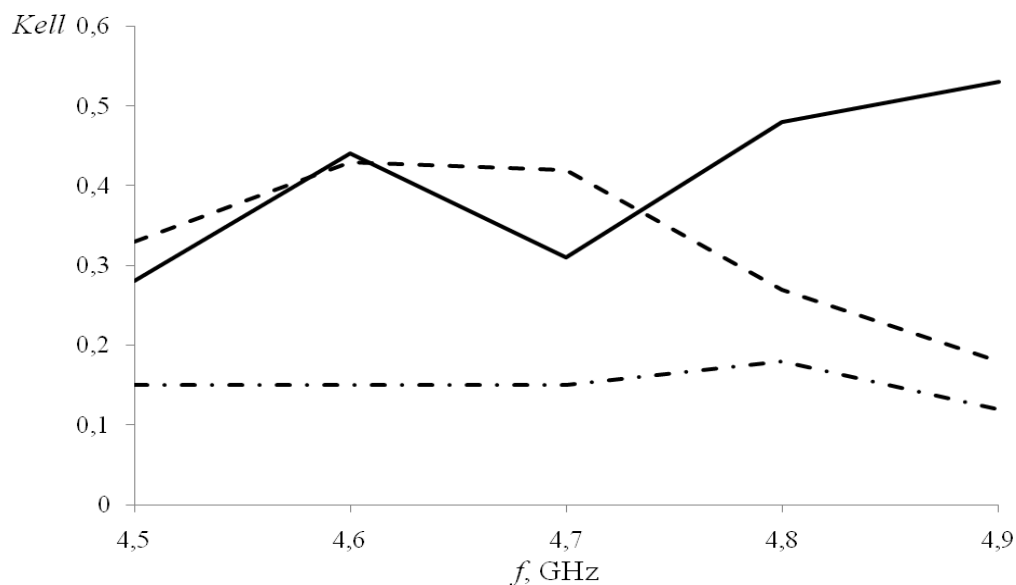


Рис. 3. Коэффициент эллиптичности в зависимости от частоты для III

Видно, что на первой поляризации максимум коэффициента эллиптичности углекомпонитных антенн из ткани сдвинут относительно максимума коэффициента эллиптичности антенны – аналога в сторону более низких частот. Для углекомпонитной антенны из нити коэффициент эллиптичности вообще не зависит от частоты. На второй поляризации величина коэффициента эллиптичности для всех антенн также практически не зависит от частоты, но параметры только углекомпонитной антенны из ткани близки к параметрам металлического аналога.

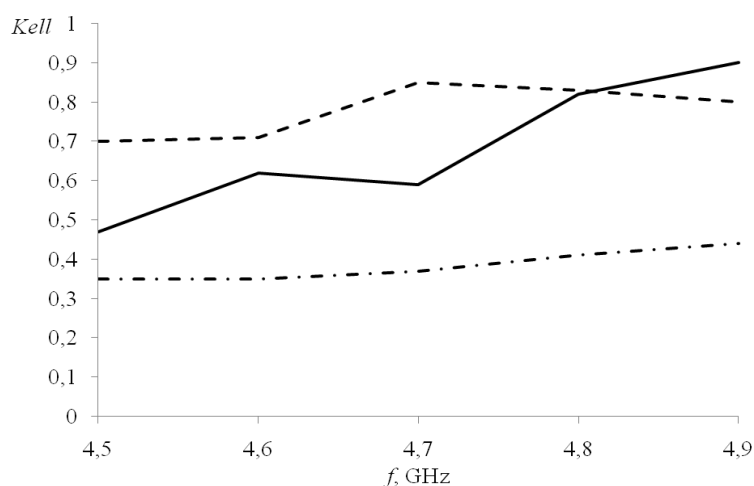


Рис. 4. Коэффициент эллиптичности в зависимости от частоты для II

Таким образом, поляризационные характеристики рупорных антенн в диапазоне 5 ГГц, выполненных из материалов с заведомо изотропной проводимостью (таких как металл и углекомпонитная ткань), близки как качественно, так и количественно. Определенные различия вызваны, видимо, небольшими конструктивными особенностями, которые были неизбежны в лабораторных технологических условиях. Применение

промышленных технологий изготовления углекомпонитных устройств должно свести к минимуму различия в их радиочастотных параметрах.

Поляризационные характеристики рупорных антенн на частотах 1,6 и 5 ГГц, выполненных из материала с возможной анизотропией проводимости стенок волноводов (углекомпонитная нить с графеносодержащим эпоксисвязующим веществом), имеют заметные отличия от аналогичных характеристик антенн со стенками с изотропной проводимостью. В данном эксперименте это привело к искажению параметров стандартного поляризационного устройства - отсутствию или изменению процесса вращения фазы распространяющейся в волноводе волны.

Работа выполнена при поддержке Министерства науки и высшего образования РФ (проект № 0729-2020-0057 в рамках базовой части Государственного задания).

Библиографический список

1. Nikolai A. Dugin, Tatiana M. Zaboronkova, Catherine Krafft, Grigori R. Belyaev, Carbon-based Composite Microwave Annennas (Review). MDPI. Electronics, 2020, V. 9, P. 590-1—590-17., dx.doi.org/10.3390/electronics9040590.
2. Характеристики углекомпонитного материала Panex 35 [Электронный ресурс]. — Режим доступа: URL: <https://zoltek.com/products/px35/> (дата обращения: 20.01.2020).
3. Электродинамика и распространение радиоволн: учебное пособие / Неганов В.А., Осипов О.В., Раевский С.Б. [и др.]; под ред. В.А. Неганова и С.Б. Раевского. Изд. 3-е, доп. и перераб. – М.: Радиотехника, 2007. – 744с.

N.A. Dugin, G.R. Belyaev

THE CARBON MATERIALS STRUCTURE INFLUENCE ON THE HIGH FREQUENCY RANGE HORN ANTENNA PARAMETERS INVESTIGATION

Volga state university of water transport

The carbon composite material structure influence on the high-frequency antenna devices parameters is considered in this work. The horn antenna walls material conductivity influence on its gain is investigated by computer simulation. The presence of anisotropic properties of the conductivity of a composite material has been studied experimentally, depending on the method of its manufacture. For this purpose, horn antennas of the L- and C- frequency range were created from carbon composite materials and graphene-containing epoxy binder with various structures. The investigation of the polarization characteristics of these antennas indicates the presence of anisotropic conductivity of the walls of the waveguide part of antennas made of carbon composite materials, which leads to a change in their parameters.

Keywords: horn antenna, carbon composite material, graphene, polarization.

МОДЕЛИРОВАНИЕ РАСПРОСТРАНЕНИЯ ИЗЛУЧЕНИЯ ИСТОЧНИКА СВЕРХШИРОКОПОЛОСНЫХ ЭЛЕКТРОМАГНИТНЫХ ИМПУЛЬСОВ С ФОКУСИРУЮЩЕЙ АНТЕННОЙ СИСТЕМОЙ

Филиал федерального государственного унитарного предприятия «Российский федеральный ядерный центр – Всероссийский научно-исследовательский институт экспериментальной физики» «Научно-исследовательский институт измерительных систем им. Ю.Е. Седакова», Россия, г. Нижний Новгород

Настоящая работа посвящена расчетно-теоретическому моделированию характеристик антенной системы, состоящей из облучателя в виде источника сверхкоротких импульсов и офсетного параболического рефлектора. Моделирование осуществляется при помощи САПР CST Microwave Studio.

Ключевые слова: напряжение излучения, система автоматизированного проектирования, рефлекторная антенна, спектральная плотность напряженности поля, диаграмма направленности.

Введение

При решении ряда прикладных задач, связанных с формированием мощных сверхкоротких электромагнитных импульсов в пространстве, возникает необходимость численной оценки параметров сигнала в дальней зоне фокусирующей антенной системы. Для решения такого рода задач требуется разработка сравнительно сложного программного обеспечения, реализующего расчетные алгоритмы с применением численных методов [1-2] решения уравнений Максвелла. Наиболее известные системы автоматизированного проектирования, такие как CST Microwave Studio [3] и Ansoft HFSS [4], несмотря на большое количество применяемых численных методов и встроенных алгоритмов расчета, не позволяют производить оценку пространственно-временных параметров сверхкоротких (сверхширокополосных) электромагнитных импульсов.

В рамках данной работы для оценки параметров сверхкороткого электромагнитного импульса в дальней зоне излучающей системы был использован компромиссный подход: моделирование было проведено с помощью САПР с предварительным численно-аналитическим преобразованием исходных пространственно-временных характеристик импульса на апертуре излучателя к виду, пригодному для использования в САПР.

Постановка задачи

На рис. 1 представлен внешний вид расчетной модели антенной системы, состоящей из рефлекторной (фокусирующей) антенны (1 – на рис. 1) и импульсного апертурного облучателя (2 – на рис. 1).

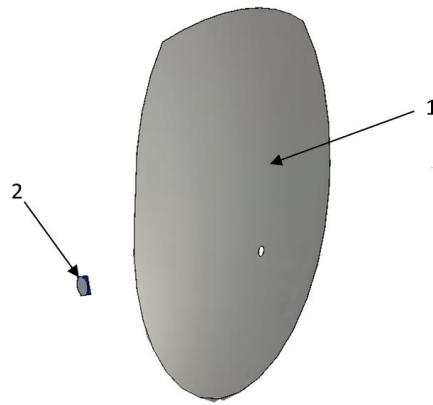


Рис. 1. Внешний вид расчетной модели антенной системы

В качестве модели фокусирующей антенны в настоящей работе была использована модель офсетной параболической антенны СТВ-3,7 марки SUPRAL с размером рефлектора $3,7 \times 4,05$ м и фокусным расстоянием 2000 мм.

В фокусе рефлекторной антенны размещена модель облучателя с диаметром апертуры 200 мм. В качестве исходных данных для расчета приведены временные зависимости напряженности электрического и магнитного полей импульсного сигнала на поверхности апертуры, а также временная зависимость напряженности поля в дальней зоне импульсного облучателя (при отсутствии рефлектора).

В рамках данной работы необходимо провести расчетно-теоретическое моделирование влияния рефлекторной параболической антенны на распространение сверхкороткого импульсного сигнала, создаваемого облучателем, а также осуществить верификацию полученных результатов.

Расчет характеристик импульсного облучателя

На рис. 2 приведена временная зависимость напряженности электрического поля в дальней зоне (на расстоянии 100 м), создаваемого импульсным облучателем в направлении максимума диаграммы направленности (ДН). Из рисунка видно, что максимальное значение напряженности электрического поля на расстоянии 100 м от модели облучателя составляет 7,1 кВ/м, что соответствует напряжению излучения $FOM = 0,71$ МВ.

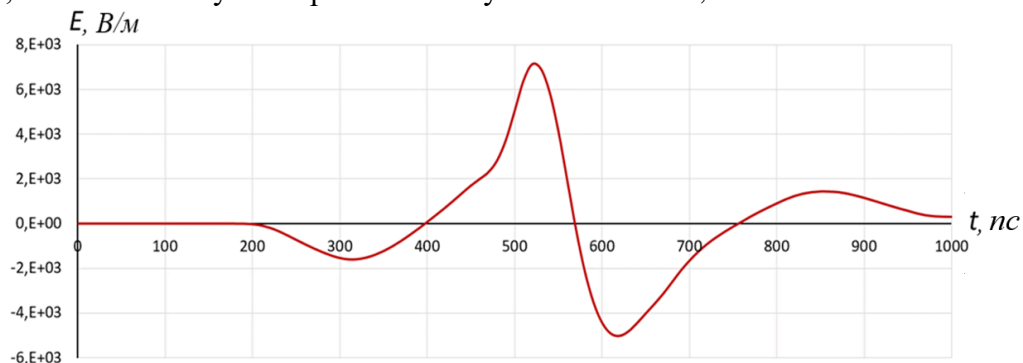


Рис. 2. Зависимость напряженности электрического поля модели импульсного облучателя от времени в дальней зоне

На рис. 3 представлена частотная зависимость спектральной плотности напряженности электрического поля, полученная в результате дискретного преобразования Фурье от временной зависимости напряженности, приведенной на рис. 2. Из рисунка видно, что значение спектральной плотности напряженности поля уменьшается более, чем в 2 раза на частотах ниже 1,5 ГГц и выше 4,5 ГГц. При этом ее максимум находится в окрестности частоты, равной 3 ГГц. Следовательно, далее будем рассматривать только указанный диапазон частот.

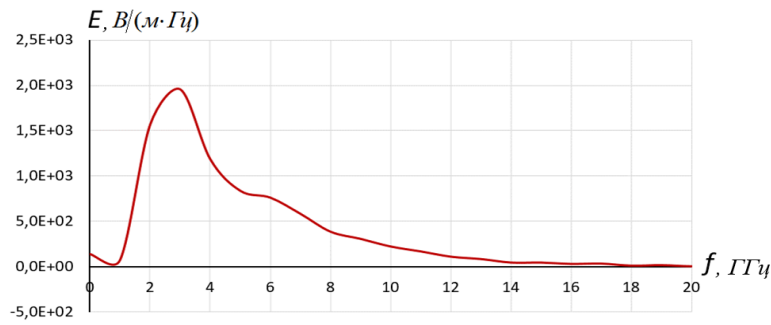


Рис. 3. Зависимость спектральной плотности напряженности электрического поля импульсного облучателя от частоты в дальней зоне

Временные зависимости компонент электрического и магнитного полей на поверхности апертуры импульсного облучателя приведены на рис. 4-6.

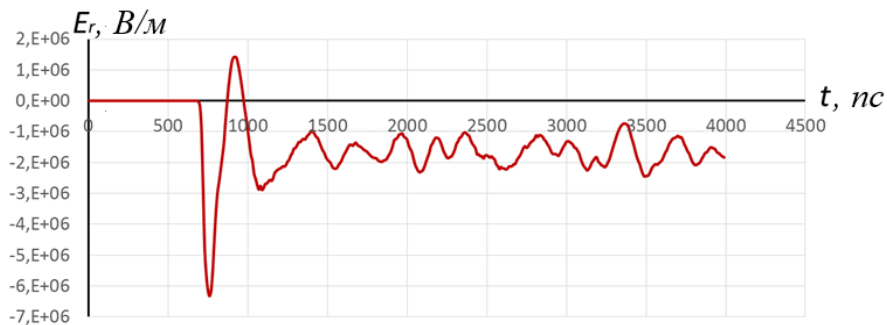


Рис. 4. Временная зависимость компоненты поля E_r на поверхности облучателя



Рис. 5. Временная зависимость компоненты поля H_ϕ на поверхности облучателя

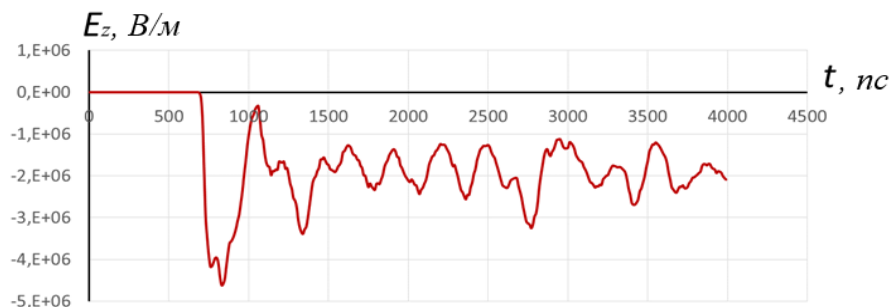


Рис. 6. Временная зависимость компоненты поля E_z на поверхности облучателя

На рис. 7-9 представлены результаты расчета частотной зависимости спектральной плотности напряженности компонент электромагнитного поля на поверхности апертуры импульсного облучателя. Для использования САПР CST Microwave Studio при расчете характеристик макета излучающего модуля и антенной системы, в состав которой входит параболический рефлектор, необходимо получить частотные зависимости спектральных составляющих электромагнитного импульса на поверхности апертуры облучателя в

декартовой системе координат. С этой целью было выполнено дискретное преобразование Фурье исходных данных при помощи численных алгоритмов, встроенных в программный пакет Matlab R2015a, а также преобразование компонент полей при переходе из цилиндрической системы координат в декартову.

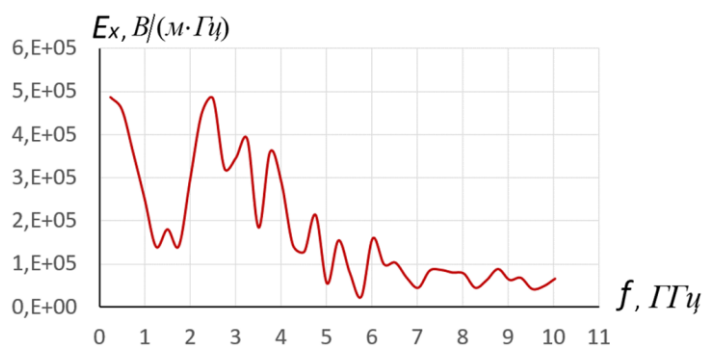


Рис. 7. Частотная зависимость спектральной плотности напряженности E_x на поверхности апертуры импульсного облучателя

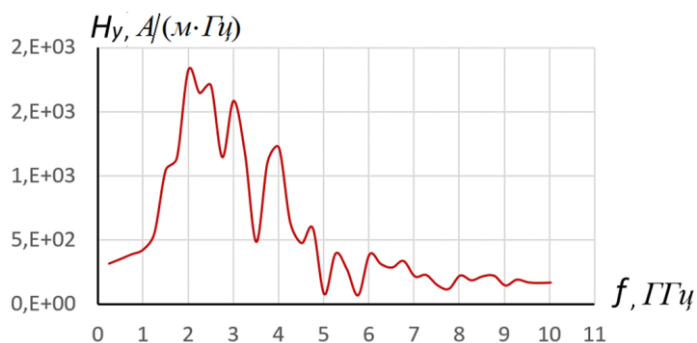


Рис. 8. Частотная зависимость спектральной плотности напряженности H_y на поверхности апертуры импульсного облучателя

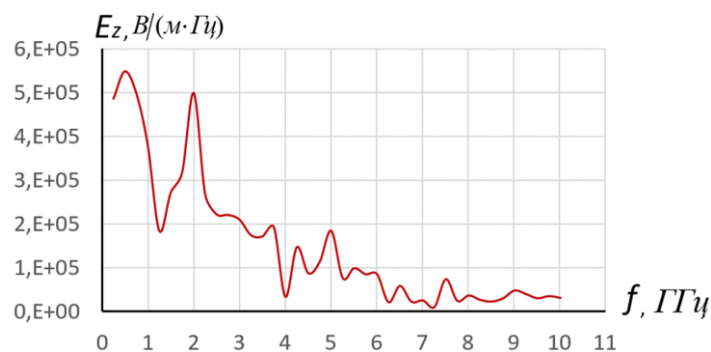


Рис. 9. Частотная зависимость спектральной плотности напряженности E_z на поверхности апертуры импульсного облучателя

Полученные результаты расчета были использованы в качестве источника в виде поля ближней зоны (nearfield source) модели импульсного облучателя в CST Microwave Studio при моделировании поля в дальней зоне макета антенной системы. На рис. 10 представлены распределения напряженностей электрического и магнитного полей на поверхности апертуры модели импульсного облучателя при фиксированном значении фазы.

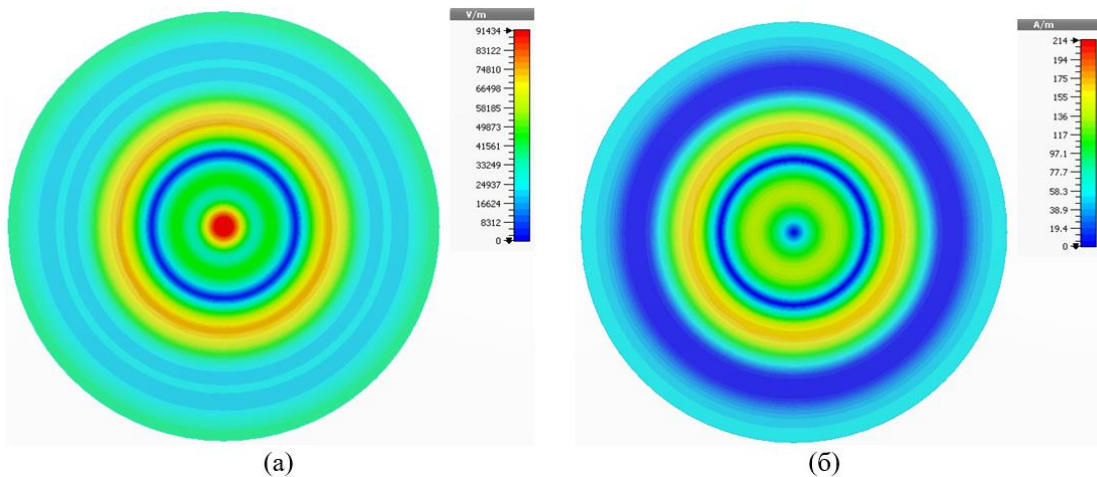


Рис. 10. Распределение электрического (а) и магнитного (б) полей на поверхности апертуры облучателя

Результаты расчета характеристик антенной системы в дальней зоне

На рис. 11-12 приведены ДН модели импульсного облучателя, на рис. 13-14 – ДН модели антенной системы при различных частотах.

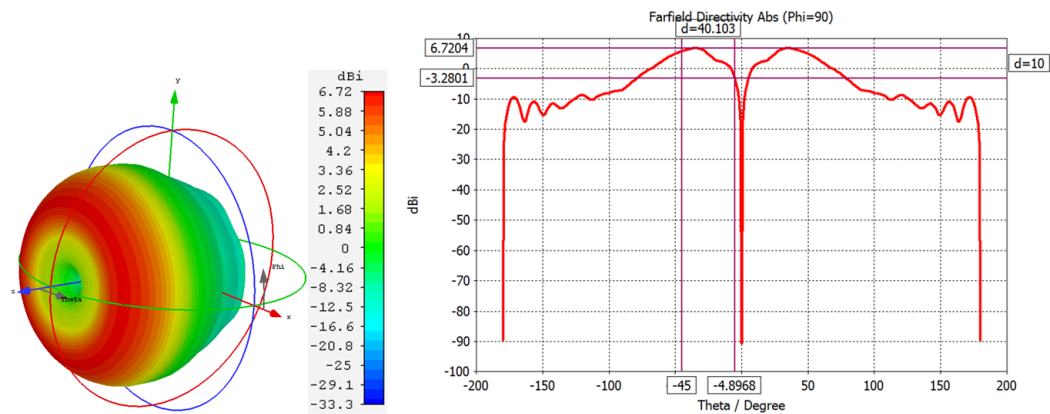


Рис. 11. ДН модели импульсного облучателя на частоте $f = 3$ ГГц

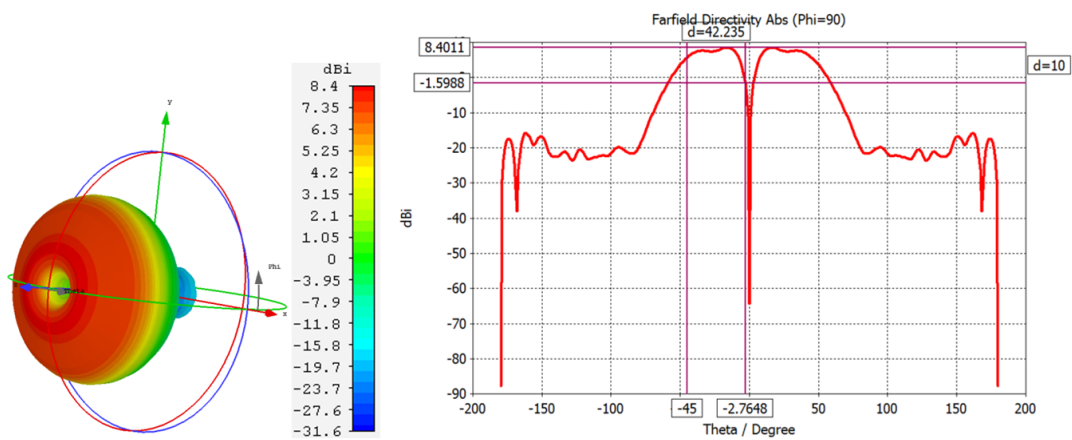


Рис. 12. ДН модели импульсного облучателя на частоте $f = 4$ ГГц

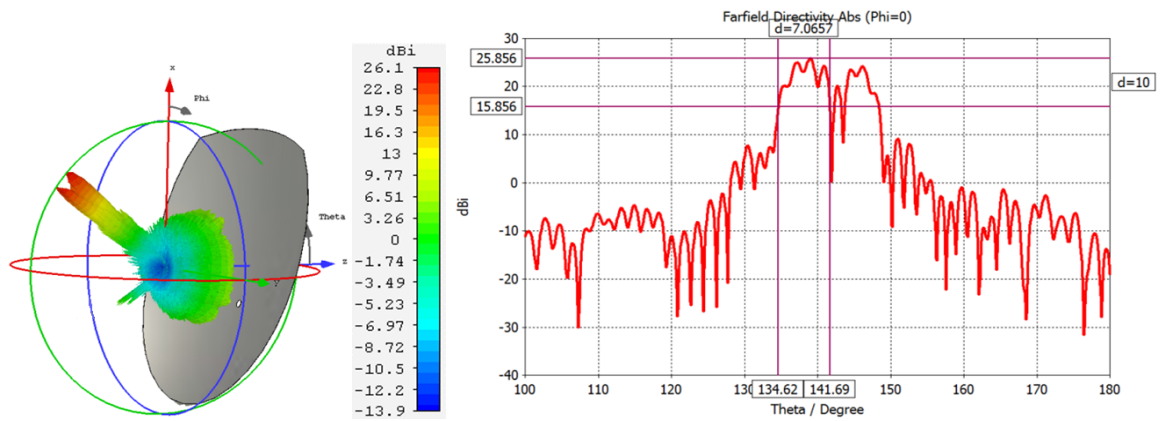


Рис. 13. ДН модели антенной системы на частоте $f = 3$ ГГц

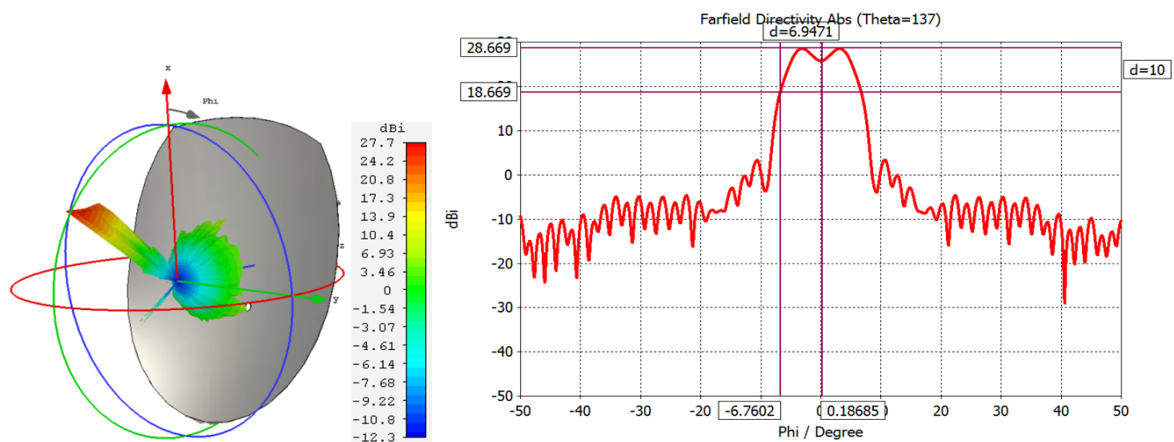


Рис. 14. ДН модели антенной системы на частоте $f = 4$ ГГц

Анализ результатов численного моделирования показывает, что параболический рефлектор обеспечивает значительное сужение (на порядок) ДН рассматриваемой модели импульсного облучателя.

Значения величин спектральной плотности напряженности электрического поля в дальней зоне (на расстоянии 100 м в направлении максимума ДН), создаваемые макетом импульсного облучателя ($E_{ИО}$) и расчетной моделью антенной системы ($E_{АС}$), а также значения усиления спектральной плотности напряженности электрического поля макета импульсного облучателя, полученного за счет применения рефлектора приведены в таблице 1. Из указанной таблицы видно, что в рассматриваемом диапазоне частот (1.5-4.5 ГГц) параболический рефлектор обеспечивает усиление спектральной плотности напряженности поля не менее, чем на 23 дБ.

Таблица 1

Частота, f (ГГц)	$E_{ИО}$ ($B/(м \cdot Гц)$)	$E_{АС}$ ($B/(м \cdot Гц)$)	Величина усиления, дБ
1.5	190	2560	23
2.0	530	8500	24
2.5	680	10000	23
3.0	380	7330	26
3.5	200	3830	25
4.0	360	6260	25
4.5	300	5400	25

Среднее по рассматриваемому диапазону частот значение усиления спектральной плотности напряженности поля за счет применения рефлектора составляет 24,4 дБ.

С целью оценки корректности полученных результатов моделирования было проведено экспериментальное исследование образца излучателя сверхширокополосных импульсов с характеристиками, близкими к характеристикам модели импульсного облучателя, рассмотренного в настоящей работе. Напряжение излучения образца излучателя в направлении максимума ДН составило: $FOM = 0,8$ МВ. Кроме того, была проведена экспериментальная оценка коэффициента усиления офсетной параболической антенны СТВ-3,7 марки SUPRAL, в фокусе которой был размещен указанный образец излучателя. Напряжение излучения антенной системы (рефлектор и образец облучателя) в направлении максимума ДН составило: $FOM = 14$ МВ. Таким образом, рефлекторная антенна позволяет увеличить амплитуду напряженности поля сверхкороткого импульса на 24,9 дБ.

Выводы

В рамках настоящей работы проведено расчетно-теоретическое моделирование излучения сверхкороткого электромагнитного импульса антенной системой, включающей модель облучателя с заданными параметрами поля на поверхности апертуры и модель офсетного параболического рефлектора. Поставленная задача была решена с использованием САПР CST Microwave Studio и среды Matlab R2015a. Для моделирования с использованием САПР исходные пространственно-временные характеристики облучателя предварительно были преобразованы в частотные зависимости спектральной плотности напряженности поля путем применения численных алгоритмов, включающих дискретное преобразование Фурье.

С целью проверки корректности полученных результатов моделирования параметров антенной системы в рамках настоящей работы было проведено экспериментальное исследование, показавшее соответствие теоретических и экспериментальных результатов.

Библиографический список

1. Марчук Г. И. Введение в проекционно-сеточные методы / Г. И. Марчук, В. И. Агошков – М.: Мир, 1979. – 392 с.
2. Сегерлинд Л. Применение метода конечных элементов / Л. Сегерлинд – М.: Мир, 1979. – 351 с.
3. Курушин А. А. Проектирование СВЧ устройств в среде CST Microwave Studio / А. А. Курушин, А. Н. Пластиков – М.: Издательство МЭИ, 2011. – 155 с.
4. Банков С. Е. Анализ и оптимизация СВЧ структур с помощью HFSS / С. Е. Банков, А. А. Курушин, В. Д. Разевиг – М.: Солон-Пресс, 2004. – 283 с.

A.S. Belov, S.M. Garanin

SIMULATION OF RADIATION PROPAGATION OF A SOURCE OF ULTRA-WIDEBAND ELECTROMAGNETIC PULSES WITH FOCUSING ANTENNA SYSTEM

Branch of the Federal State Unitary Enterprise «Russian Federal Nuclear Center - All-Russian Scientific Research Institute of Experimental Physics» «Scientific Research Institute of Measuring Systems named after Yu.E. Sedakov», Russian Federation, Nizhny Novgorod

The article is devoted to calculation-theoretical modeling of characteristics of antenna system consisting of irradiator in the form of source of ultra-short pulses and offset parabolic reflector. Simulation is carried out using CAD CST Microwave Studio.

Keywords: radiation voltage, computer-aided engineering system, reflector antenna, spectral density of field strength, directional pattern.

МАЛОГАБАРИТНЫЙ ПОЛОСНО-ПРОСКАЮЩИЙ ФИЛЬТР С ПОВЫШЕННЫМ УРОВНЕМ ПРОХОДЯЩЕЙ СРЕДНЕЙ МОЩНОСТИ

Филиал Федерального государственного унитарного предприятия "Российский федеральный ядерный центр - Всероссийский научно-исследовательский институт экспериментальной физики" "Научно-исследовательский институт измерительных систем им. Ю.Е. Седакова"

В представленной работе предложен вариант построения полосно-пропускающего фильтра (ППФ) на основе коаксиальных керамических линий круглого сечения, обеспечивающий повышенный уровень проходящей средней мощности. Представленная конструкция ППФ может быть использована в передающих устройствах для выходной фильтрации СВЧ-сигналов с целью исключения внеполосных и побочных излучений или для входной предварительной фильтрации полезного сигнала в приёмниках, работающих в условиях воздействия мощных активных помех.

Ключевые слова: коаксиальный керамический резонатор, фильтр СВЧ, полосно-пропускающий фильтр, САПР.

Введение

В настоящее время в классе устройств СВЧ диапазона есть большая линейка фильтров, обеспечивающих характеристики на высоком уровне. Конструкции ППФ на встречных стержнях и на дисковых диэлектрических резонаторах, выполненных из керамического материала, обеспечивают широкий диапазон полос пропускания и повышенный уровень импульсной мощности, но имеют низкие массогабаритные характеристики и низкую стойкость к механическим воздействиям [1,2]. Успешно применяются СВЧ ППФ на четвертьволновых коаксиальных керамических резонаторах квадратного сечения [3,4] обеспечивающие высокие значения характеристик ППФ при малых габаритах устройства. Предлагаемая в представленной работе конструкция обладает рядом преимуществ, позволяющих конкурировать с аналогами по массогабаритным, электрическим и тепловым показателям.

Описание конструкции

Внешний вид фильтра представлен на рис. 1 - 2. Предлагаемый СВЧ ППФ состоит из гальванически соединённых между собой боковыми поверхностями четвертьволновых резонаторов, изготовленных на основе коаксиальных керамических линий (рис.1).

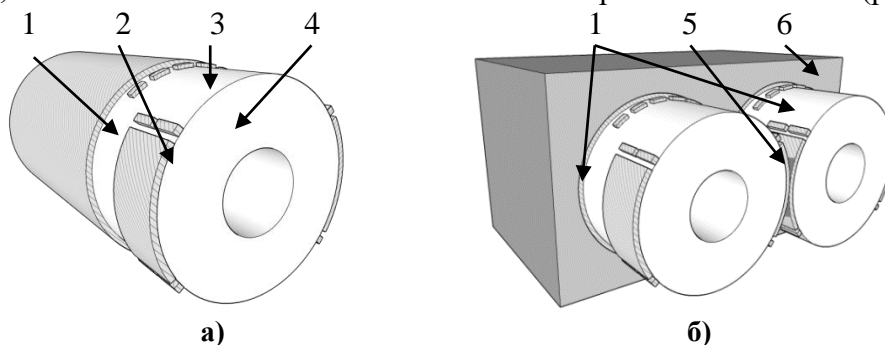


Рис. 1. Конструкция ППФ без установки диэлектрических стержней и основания до (а) и после (б) установки в корпус: 1 – коаксиальный керамический резонатор круглого сечения; 2 – зазор в металлизации внешнего проводника; 3 – дополнительные контактные площадки для регулировки частоты; 4 – дополнительные контактные площадки для регулировки ёмкости связи; 5 – место пайки; 6 – корпус ППФ

Выбор круглого сечения для резонаторов позволяет более качественно проводить их соединение и исключить избыточный нагрев на ребрах, как это было бы при использовании резонаторов квадратного сечения.

Существенным преимуществом представленной конструкции является то, что каждый резонатор и ёмкости связи изготовлены как один конструктивный элемент. Ёмкости связи выполнены в виде керамических конденсаторов, обкладками которых являются внутренние проводники и внешние поверхности отрезка коаксиальной керамической линии. Ёмкости отделены от резонатора зазором в металлизации внешней поверхности коаксиальной керамической линии.

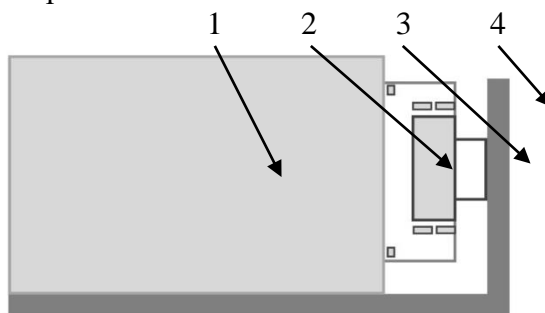


Рис. 2. Боковой вид конструкции ППФ с основанием

1 – корпус ППФ; 2 – коаксиальный керамический резонатор круглого сечения; 3 – диэлектрические стержни круглого сечения с металлизацией на торцах; 4 – металлическое основание

При прохождении высокого уровня мощности через фильтр, значительным фактором становятся выделяемая тепловая энергия. Электротепловой пробой обусловлен нагревом диэлектрика, связан с условиями охлаждения материала, температурой окружающей среды и с нагревостойкостью материала [5]. Напряжение пробоя определяется по формуле:

$$U_{np} = k \sqrt{\frac{Sh}{fa\epsilon tg\delta}}, \quad (1)$$

где k - числовой коэффициент, равный $1,15 \cdot 10^5$;

S - коэффициент теплопередачи системы диэлектрик - металл электродов;

h - толщина диэлектрика;

f - частота;

a - температурный коэффициент тангенса угла диэлектрических потерь;

ϵ - диэлектрическая проницаемость;

$tg\delta$ - тангенс угла потерь диэлектрика при t_0 - температуре окружающей среды.

Область применения фильтра может значительно расширится при обеспечении отвода тепла от резонансной части фильтра. Один из предлагаемых способов – помещение фильтра в корпус. Корпус следует изготавливать из металла с высоким значением коэффициента теплопроводности в форме параллелепипеда (устойчивая форма) с глухими отверстиями круглой формы для установки туда резонаторов. В центре каждого отверстия устанавливается штырь круглого сечения с диаметром, равным диаметру внутреннего проводника резонатора и с длиной равной длине коаксиальных керамических линий. Установка штыря в центре коаксиального керамического резонатора позволяет реализовать тепловой контакт корпуса с внутренним проводником резонатора, где наблюдается максимальное значение температуры [6]. Размещение в корпус обеспечивает повышенные прочностные характеристики, защищая место соединения резонаторов при ударах и вибрациях. Дополнительно прочность обеспечивается гальваническим соединением резонаторов с корпусом.

При таком конструктивном решении тепловая энергия отводится от внутреннего проводника резонатора в окружающее пространство. Открытый конец может быть дополнительно соединен с диэлектрическим стержнем (рис. 2), изготовленным из материала с высоким коэффициентом теплопроводности, также позволяя отвести тепловую энергию и исключить градиент температур по длине ППФ. Длина диэлектрического стержня может быть определена из величины напряжения теплового пробоя и исходя из требований к габаритам ППФ. Пробивное напряжение будет выше, если диэлектрик будет толще, условия теплоотвода лучше, частота ниже, а диэлектрическая проницаемость и тангенс угла диэлектрических потерь меньше. Для крепления диэлектрических стержней их торцы металлизируются, один конец соединяется с металлическим стержнем корпуса, а второй к вертикальной поверхности основания выполненного в виде уголка (рис. 2). При установке ППФ автономно, когда тепло с поверхности корпуса рассеивается в окружающее пространство без вреда радиоэлектронной аппаратуре, на боковые поверхности корпуса могут быть установлены радиаторы.

Процесс сборки ППФ

Помимо указанных выше преимуществ, предлагаемая конструкция ППФ технологична в сборке и регулировке. Возможный вариант последовательности сборки и регулировки предлагаемого ППФ:

1. Изготовление коаксиальных керамических линий с формированием резонансной части длиной соответствующей четверти длины волны в керамическом материале, емкостей связи резонаторов и дополнительных контактных площадок для регулировки емкостей связи и частоты.

2. Соединение пайкой боковых поверхностей соседних резонаторов и конденсаторов связи между резонаторами.

3. Изготовление корпуса из металла с высоким значением коэффициента теплопроводности. Корпус выполнен в форме параллелепипеда с глухими отверстиями круглой формы с штырями круглого сечения в центре каждого из отверстия. Диаметр отверстия равен диаметру внешнего проводника резонатора. Диаметр штырей равен диаметру внутреннего проводника резонатора, длина штырей равна длине коаксиальных керамических линий.

4. Установка резонаторов в корпус.

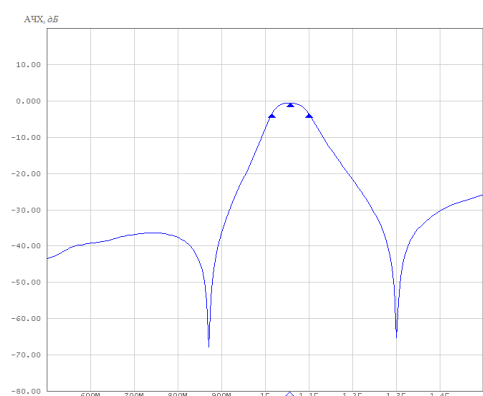
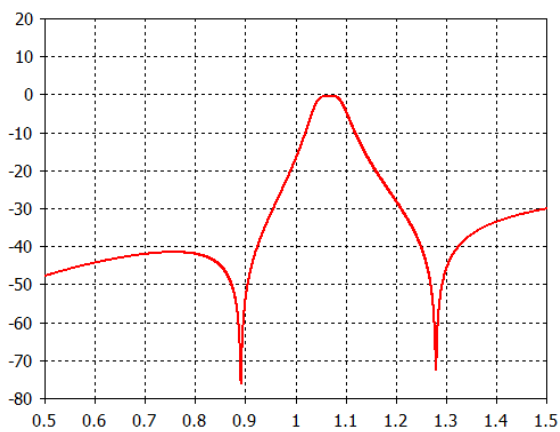
5. Установка встык к металлическим штырям корпуса диэлектрических стержней, изготовленных из материала с высоким коэффициентом теплопроводности.

6. Установка резонаторов в корпусе на основание, выполненное в виде уголка. Установка на горизонтальную поверхность уголка корпуса, прикрепление к вертикальной поверхности основания торцевых частей диэлектрических стержней.

7. Установка радиаторов на боковые части корпуса при необходимости дополнительного отвода тепла в изделии в окружающее пространство.

Характеристики ППФ с повышенным уровнем проходящей средней мощности

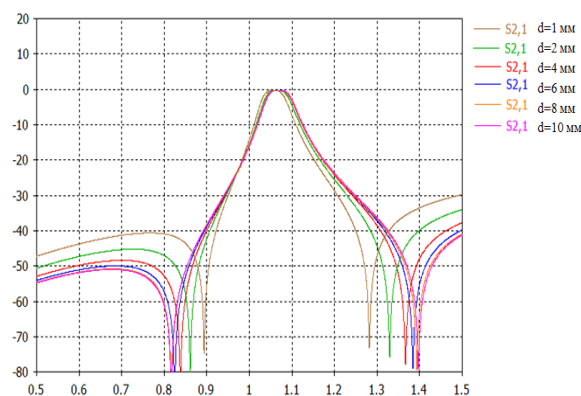
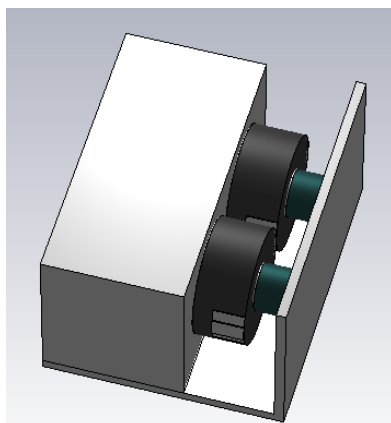
Необходимость применения в конструкции фильтра на коаксиальных резонаторах дополнительных отводящих тепло элементов обусловлено выделяемой мощностью, вызванной проводимостью металлического покрытия и тепловыми потерями в диэлектрике. Высокий уровень электрических характеристик ППФ подтверждается электромагнитным анализом в системе автоматического проектирования (САПР). На рис. 3 представлены характеристики, рассчитанные для модели в САПР и измеренные АЧХ ППФ. Анализируемая структура состоит из двух ККР, соединенных боковыми поверхностями и установленными в металлический корпус.



а) б)
Рис. 3. Рассчитанная для модели (а) и измеренная (б) АЧХ ППФ

Расчеты проводились для ППФ построенного на основе ККР (диэлектрическая проницаемость 100) с диаметрами внутреннего проводника 3 мм и внешнего проводника 6 мм. При центральной частоте ~ 1 ГГц и ширине полосы пропускания 80 МГц КСВн на центральной частоте составляет 1,07.

На рис. 4 представлена модель ППФ СВЧ при добавлении диэлектрических стержней и основания для улучшения теплоотвода, и обеспечения повышенной средней мощности СВЧ, а также рассчитанные характеристики модели при различных длинах диэлектрических стержней.



а) б)

Рис. 4. Внешний вид (а) и АЧХ при разной длине диэлектрических стержней (б) модели ППФ СВЧ

Добавление основания и стержней с диэлектрической проницаемостью равной 10 и малой длиной не приводит к значительным изменениям АЧХ. При увеличении длины диэлектрических стержней происходит увеличение полосы пропускания фильтра со смещением частот провалов. Использование для диэлектрических стержней материалов с большей диэлектрической проницаемостью приводит к смещению центральной частоты фильтра в сторону меньших частот и росту КСВн при почти неизменном положении провалов АЧХ.

Выводы

Создание ППФ на ККР с повышенным уровнем пропускаемой СВЧ-мощности является актуальной задачей. Предлагаемая в представленной работе конструкция ППФ обладает малыми габаритами и массой, низкими потерями в полосе пропускания, а также повышенными, по сравнению с аналогами, прочностными характеристиками. Возможная область практического применения подобного ППФ - приемо-передающая аппаратура системы телеметрии, для выполнения требований обеспечения ЭМС, а также в дуплексах базовых станций сотовой связи и ретрансляторов радиорелейных линий связи.

Библиографический список

1. Адрианов А.В., Зикий А.Н. Моделирование и экспериментальное исследование узкополосного фильтра на встречных стержнях //Инженерный вестник Дона, 2018. вып. 3 – 44 с.
2. Ильченко М.Е. Диэлектрические резонаторы / М.: Радио и связь, 1989. – 224 с.
3. Седаков А.Ю. Антенны и функциональные узлы СВЧ- и КВЧ-диапазонов. Методы расчёта и технология изготовления – М.: Радиотехника, 2011. – 101 с.
4. Патент на изобретение 2619363 РФ МПК Н01Р 1/219 (2006.01). Полосно-пропускающий СВЧ-фильтр / Шишкин Д.Р., Кунилов А.Л., Балобанов Е.С., Ивойлова М.М. – № 2016129322; Заявл. 18.07.16; Оpubл. 15.05.17, Бюл. № 14.
5. Привалов Е. Е. Основы электроматериаловедения: учебное пособие – Москва, Директ-Медиа, 2017. -160 с.
6. Семёнов Н.А. Техническая электродинамика. Учебное пособие для вузов. М., «Связь», 1973. – 242 с.

E.S. Balobanov, M.M. Ivoilova, A.L. Kunilov, M.S. Nechaeva

SMALL-SIZED BANDPASS FILTER WITH INCREASED PASSED AVERAGE POWER

Branch of the Federal State Unitary Enterprise "Russian Federal Nuclear Center - All-Russian Research Institute of Experimental Physics" "Scientific Research Institute of Measuring Systems named after Yu.E. Sedakov"

Purpose: construction of a band-pass microwave filter, providing an increased level of transmitted average power

Design/methodology/approach: The proposed microwave filter consists of quarter-wave resonators galvanically connected to each other by the side surfaces, made on the basis of coaxial ceramic lines and placed in a case.

Findings: The presented filter design can be used in transmitting devices for output filtering of microwave signals in order to exclude out-of-band and spurious emissions or for input pre-filtering of a useful signal in receivers operating under the influence of powerful active interference.

Originality/value: The filter design proposed in the presented work has small dimensions and weight, low losses in the passband, as well as increased strength characteristics compared to analogs.

Keywords: coaxial ceramic resonators, microwave filter; band-pass filter, CAD.

ПОЛИРОВАНИЕ ПОДЛОЖЕК ИЗ КОРУНДОВОЙ КЕРАМИКИ ДЛЯ КОМПОНЕНТОВ СВЧ УСТРОЙСТВ

*Филиал Федерального государственного унитарного предприятия «Российский федеральный ядерный центр – Всероссийский научно-исследовательский институт экспериментальной физики» «Научно-исследовательский институт измерительных систем им. Ю.Е. Седакова»
г. Нижний Новгород*

Исследован процесс механического полирования подложек из корундовой керамики с использованием полировальных тканей. Представлены результаты обработки подложек с применением полировальных тканей с различными свойствами и полирующих суспензий с разными зернистостями абразива.

Ключевые слова: полирование; корундовая керамика; алмазная суспензия; полировальные ткани.

Эффективность технологического обеспечения проектирования и изготовления компонентов бортовых радиоэлектронных устройств [1] во многом зависит от имеющейся в распоряжении разработчиков технологической базы [2], формирование и совершенствование которой является непрерывным процессом и осуществляется одновременно с проектированием и изготовлением устройств [1].

Одним из материалов, используемых в разработках и изготовлении современных компонентов СВЧ устройств, является радиотехническая керамика. Благодаря своим механическим, теплопроводным и электрическим характеристикам керамика применяется в качестве конструкционной основы компонентов бортовых сверхвысокочастотных (СВЧ) устройств [3], например, в качестве подложек, изготавливаемых СВЧ микроплат [4], изоляторов для резонаторов микрополосковых антенн бортовой аппаратуры спутниковой навигации [5] и др.

Целью работы являлось исследование и оптимизация технологии полирования пластин из корундовой керамики «Кортим» [5] (собственной разработки и изготовления), применяемой в конструкциях СВЧ устройств, разрабатываемых на предприятии, с использованием технологической базы «НИИИС им. Ю.Е. Седакова».

В процессе работы было необходимо исследовать качество обрабатываемых поверхностей полированием подложек из керамики «Кортим» с применением полировальных тканей в качестве материала полировальника. Ожидалось, что применение полировальных тканей может позволить уменьшить размеры и концентрацию поверхностных дефектов обрабатываемых поверхностей подложек. К тому же, возможность быстрой смены на полировальнике полировальных тканей в процессе обработки при переходе от алмазной суспензии одной зернистости к суспензии другой зернистости могло сократить производственный цикл обработки.

Требования по точности выполнения размеров и формы, а также качество обработанных поверхностей, могут быть различными в зависимости от назначения детали. Так, например, для подложек, используемых для изготовления СВЧ компонентов с микрополосковыми элементами, качество обработанных поверхностей должно обеспечивать хорошую адгезию с формируемыми на ней полосковыми или пленочными проводниками. В зависимости от технологии формирования проводников типовые требования – Ra менее 0,32 мкм.

Операция полирования является заключительной операцией, необходимой для доводки обрабатываемой поверхности с целью достижения требуемой шероховатости и точности формы поверхности. Из всей последовательности операций механической обработки полирование является наиболее длительным и трудоемким процессом. На качество получаемой полированной поверхности влияют многие взаимозависимые факторы, такие как продолжительность полирования, давление, материал полировальника и полирующие абразивы.

В качестве объекта обработки применялись экспериментальные образцы из корундовой керамики «Кортим» в виде подложек 2x32x32 мм, предварительно обработанные (шлифованные по габаритам и толщине в заданные размеры).

Подложки крепились на планшайбу с помощью специальной наклеочной смолы. В качестве материала полировальника использовались полировальные ткани SAMBA-N, MAMBO, ALUPOL-PLUS производства компании Met Cata GmbH, Германия. Применение полировальных тканей указанной компании было обусловлено доступностью к началу проведения работ.

Краткое описание применяемых в работе полировальных тканей:

- SAMBA-N – плотный тканый материал средней твердости из натурального хлопка диаметром 250 мм, самоклеющийся;
- MAMBO – полотно, покрытое слоем пористой резины, диаметром 250 мм, самоклеющееся;
- ALUPOL-PLUS – синтетическая замшевая ткань диаметром 250 мм, самоклеющаяся.

Одно из преимуществ применения самоклеющихся полировальных тканей является то, что их можно наклеивать на любые металлические или стеклянные шлифовальники, а также быстро производить их замену в производственном цикле изготовления деталей.

В качестве абразива применялись моно- и поликристаллические алмазные суспензии MAGNUM-TOP-DUO и POLY-TOP-DUO на водной основе (производства той же компании – Met Cata GmbH). Данные суспензии уже готовы к применению. Подача суспензий на полировальную поверхность осуществлялась вручную, с помощью кисти. Низкая вязкость суспензий позволяет быстро смачивать полировальную ткань и равномерно распределять ее по полировальной поверхности.

Полировальные ткани наклеивались на типовой стеклянный шлифовальник (полировальник).

Полирование подложек производилось на шлифовально-полировальном станке модели ЗШП-350М и проводилось в следующей последовательности:

1. На полировальнике с тканью для полирования SAMBA-N с применением монокристаллической алмазной суспензии MAGNUM-TOP-DUO зернистостью 9 мкм.
2. На полировальнике с тканью для полирования MAMBO с применением монокристаллической алмазной суспензии MAGNUM-TOP-DUO зернистостью 3 мкм.
3. На полировальнике с тканью для полирования ALUPOL-PLUS с применением поликристаллической алмазной суспензии POLY-TOP-DUO зернистостью 3 мкм.
4. На полировальнике с тканью для полирования MAMBO с применением монокристаллической алмазной суспензии MAGNUM-TOP-DUO зернистостью 3 мкм.

В процессе полирования снимаемый припуск и плоскостность контролировали с помощью индикатора многооборотного 1 МИГ-1, закрепленного на индикаторной стойке, шероховатость обработанных поверхностей – с помощью приборов для измерения шероховатости MarSurf PS1 и HOMMEL TESTER, качество обработанной поверхности оценивали с помощью микроскопа МБС-9 и высокоточной измерительной системы μ Surf explorer (производства NanoFocus AG, Германия) по наличию и количеству дефектов (царапины, включения, поры, раковины, выколы).

Неплоскостность обработанных подложек составила от 0,005 до 0,01 мм.

Время обработки, снимаемый припуск по переходам, шероховатость обработанной поверхности на шлифовально-полировальном станке ЗШП-350М приведены в табл. 1.

Таблица 1

Результаты обработки (полирования)

Алмазная суспензия / материал полировальника	Время обработки, мин	Снимаемый припуск, мм	Шероховатость поверхности Ra, мкм	Длина трассирования L, мм
MAGNUM-TOP-DUO, 9 мкм / SAMBA-N	180	0,002 – 0,003	0,43 – 0,35	4,5
MAGNUM-TOP-DUO, 3 мкм / MAMBO	180	0,001	0,30 – 0,25	4,5
POLY-TOP-DUO, 3 мкм / ALUPOL-PLUS	100	0	0,28 – 0,23	4,5
MAGNUM-TOP-DUO, 3 мкм / MAMBO	100	0,001	0,21 – 0,19 0,16 – 0,14	4,5 1,5

При полировании монокристаллической алмазной суспензией MAGNUM-TOP-DUO зернистостями 9 мкм и 3 мкм на обработанной поверхности наблюдался алмазный фон: отдельные тонкие «исчезающие» риски, меняющие направление при изменении направления освещения поверхности.

При полировании поликристаллической алмазной суспензией POLY-TOP-DUO зернистостью 3 мкм на обработанной поверхности отмечалось появление «ряби». Поэтому продолжили полирование суспензией MAGNUM-TOP-DUO зернистостью 3 мкм на полировальнике с тканью для полирования MAMBO.

При полировании суспензией MAGNUM-TOP-DUO зернистостью 9 мкм наблюдалось появление блеска на обработанных поверхностях. При полировании суспензией MAGNUM-TOP-DUO зернистостью 3 мкм наблюдалось снижение размеров и концентрации поверхностных дефектов на обработанных поверхностях по отношению к обработанным поверхностям суспензией MAGNUM-TOP-DUO зернистостью 9 мкм.

Шероховатость обработанных поверхностей контролировалась с помощью приборов для измерения шероховатости Mar Surf PS1 и HOMMEL TESTER, но на разной длине трассирования измерения давали разные результаты.

Значения шероховатости Ra обработанных поверхностей подложек, полученные при измерении с помощью прибора Mar Surf PS1 на длине трассирования L=4,5 мм, отличаются от значений шероховатости Ra, измеренных с помощью прибора HOMMEL TESTER на длине трассирования L=1,5 мм. После полирования суспензией MAGNUM-TOP-DUO 3 мкм измеренное значение шероховатости обработанной поверхности составило Ra0,16 – 0,14 мкм на длине трассирования L=1,5 мм, а на длине трассирования L=4,5 мм измеренное значение шероховатости обработанной поверхности составило от Ra0,16 – 0,14 мкм. Можно предположить, что размеры и концентрация поверхностных дефектов хоть и малы, но на более протяженных участках трассирования (L=4,5 мм) оказывают влияние на измеренное значение шероховатости.

По данным проведенных исследований, с увеличением продолжительности полирования алмазной суспензией MAGNUM-TOP-DUO зернистостью 3 мкм снижается относительное содержание дефектов, а при времени полирования более 280 мин остается практически постоянным (см. табл. 1).

При шероховатости от Ra 0,16 до Ra 0,14 мкм на обработанных поверхностях подложек имелись поры размерами до десятков микрометров, а также царапины («алмазный фон»), оставшиеся от полировки. Из-за наличия пор и царапин измеренное значение шероховатости поверхности увеличивается до Ra 0,22 - 0,19 мкм. Указанные дефекты, как правило, не являются браковочным признаком, но в то же время

существенно снижают надежность и процент выхода годных готовых изделий (микросхем, микросборок и др.).

Применение высокоточной измерительной системы μ Surf explorer позволило получить реальное представление о морфологии поверхности и возможность точно оценить размеры дефектов. В основе высокоточной измерительной системы μ Surf explorer лежит принцип действия конфокального микроскопа [6].

На рис. 1 представлена фотография-протокол анализа исследуемой полированной поверхности пластины (увеличение объектива $20\times$, время измерения 3 с), обработанной с применением суспензии MAGNUM-TOP-DUO зернистостью 3 мкм. На фотографии отчетливо видны поры и включения, их форма и размер.

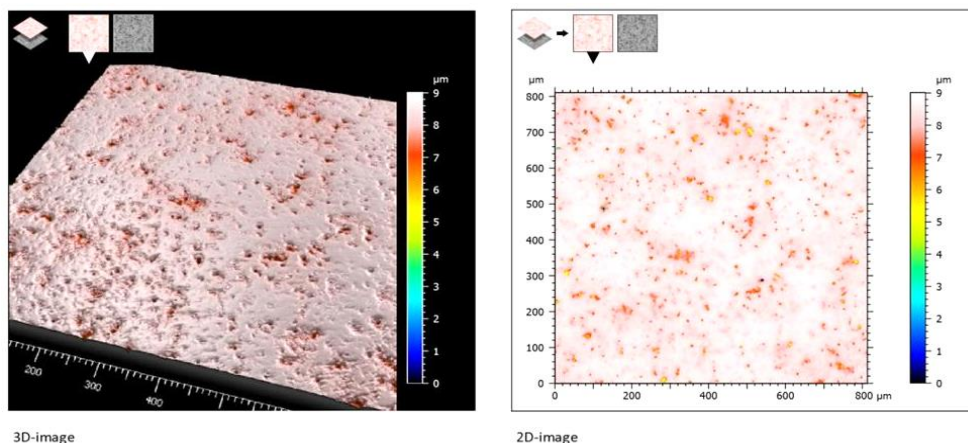


Рис. 1. Фотография-протокол анализа исследуемой полированной поверхности пластины с помощью системы μ Surf explorer

Измерение параметров шероховатости полированной поверхности с помощью μ Surf explorer дает более полное представление об исследуемой поверхности.

В соответствии с рис. 2 значение шероховатости R_a исследуемой поверхности составляет 0,150 мкм, значение шероховатости R_z составляет 1,53 мкм.

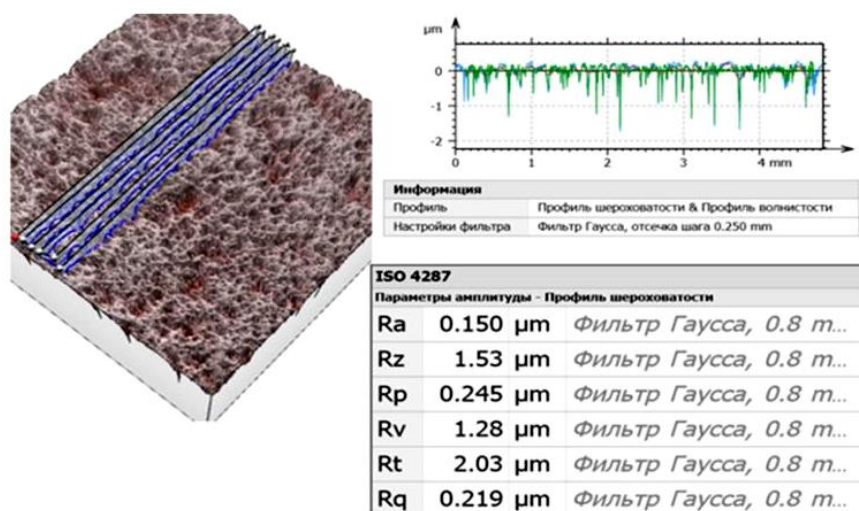


Рис. 2. Измерение параметров шероховатости полированной поверхности подложки с помощью системы μ Surf explorer

Сравнительная характеристика по измерению значений шероховатости R_a и R_z на поверхности подложки, обработанной суспензией MAGNUM-TOP-DUO зернистостью 3 мкм, с помощью различных приборов для измерения шероховатости представлена в табл. 2.

Таблица 2

Значения шероховатости поверхности после обработки суспензией 3 мкм

Прибор для измерения шероховатости, тип	Инструмент для измерения шероховатости	Длина трассирования L, мм	Значение шероховатости, мкм (класс шероховатости)	
			Ra	Rz
MarSurf PS1	Алмазная игла	4,5	0,2 (9)	1,58 (9)
HOMMEL TESTER	Алмазная игла	4,5	0,21 (9)	1,58 (9)
HOMMEL TESTER	Алмазная игла	1,5	0,144 (10)	1,54 (9)
μSurfexplorer	Бесконтактный	0,25	0,15(10)	1,53 (9)

В соответствии с данными табл. 2 значения шероховатости Ra обработанных поверхностей подложек, полученные при измерении с помощью прибора MarSurf PS1 на длине трассирования L=4,5 мм, отличаются от значений шероховатости Ra, измеренных с помощью прибора HOMMEL TESTER на длине трассирования L=1,5 мм, что подтверждает, что на более протяженных участках трассирования размеры и концентрация поверхностных дефектов (поры, включения) оказывают влияние на значение измеренной шероховатости Ra – в пределах двух классов. Значение шероховатости Rz на исследуемой поверхности остается постоянным при любой длине трассирования.

Выводы

Проведены исследования влияния параметров механического полирования на качество поверхности подложек из корундовой керамики «Кортим». Показано, что применение полировальных тканей и алмазных суспензий позволит уменьшить размеры и концентрацию поверхностных дефектов обрабатываемых поверхностей подложек.

Быстрая замена полировальных тканей в процессе цикла изготовления при переходе от алмазной суспензии порошка одной зернистости к суспензии другой зернистости позволяет сократить производственный цикл обработки.

Результаты работы позволили расширить возможности технологической базы [7] «НИИИС им. Ю.Е. Седакова» для технологического обеспечения проектирования и изготовления СВЧ устройств [1], в частности обработки подложек и пластин для компонентов СВЧ устройств.

Библиографический список

1. Козлов В.А. Развитие структуры технологического обеспечения при системном проектировании и изготовлении СВЧ компонентов бортовых радиоэлектронных систем / В.А. Козлов, Ю.А. Светлаков, А.Ю. Седаков // Проектирование и технология электронных средств. - 2018. - № 3. - С. 23-29.
2. Светлаков Ю.А. Формирование технологической базы при технологическом обеспечении проектирования и изготовления СВЧ устройств / Ю.А. Светлаков // Проектирование и технология электронных средств. – 2018. - № 4. - С. 30-36.
3. Непочатов Ю.К., Разработка корундовых и алюмонитридных подложек для изделий СВЧ техники/ Ю.К. Непочатов, А.А. Денисова, Ю.И. Щецова, В.А. Ефремов // Обмен опытом в области создания сверхширокополосных радиоэлектронных систем: материалы V Общероссийской науч.-техн. конф. -Омск: Изд-во ОмГТУ.- 2014. - С. 188-196.
5. Гудина Т.Л. Разработка технологии лазерной обработки отверстий в подложках из керамического материала ВК 100-1 с металлизацией для НЧ и СВЧ микроплат / Т.Л. Гудина, Ю.А. Светлаков// Труды XXIV Международной научно-технической конференции «Информационные системы и технологии», ИСТ-2018. – Нижний Новгород, НГТУ, 2018, с. 306-310.
6. Гудина Т.Л., Применение керамического материала «Кортим» в конструкции микрополосковой антенны бортовой аппаратуры спутниковой навигации / Т.Л.

Гудина, А.В. Кузнецов, Ю.А. Светлаков, И.П. Ярославцева // Материалы XXII Международной научно-технической конференции «Информационные системы и технологии», ИСТ-2016. – Нижний Новгород, НГТУ, 2016, С. 124.

7. Лазерная конфокальная микроскопия: метод. указания / сост. П.Е. Тимченко, Е.В. Тимченко. – Самара: Изд-во Самар. гос. аэрокосм. ун-та, 2014. – 76 с.

8. Светлаков Ю.А. Совершенствование и развитие технологической базы проектирования и изготовления СВЧ устройств / Ю.А. Светлаков // Материалы XXIII Международной научно-технической конференции «Информационные системы и технологии», ИСТ-2017. – Нижний Новгород: НГТУ им. Р.Е. Алексеева, 2017. - С. 1344-1349.

N.V. Ageycheva

THE POLISHING OF PLATES FROM CORUNDUM CERAMICS FOR COMPONENTS OF MICROWAVE DEVICES

Branch of the Federal State Unitary Enterprise “Russian Federal Nuclear Center - All-Russian Scientific Research Institute of Experimental Physics” “Scientific Research Institute of Measuring Systems named after Yu.Ye. Sedakov”, Nizhny Novgorod

The process of mechanical polishing of plates from corundum ceramics with use of fabrics for polish is investigated. Results of processing of substrates using fabrics for polish with different properties and the suspensions for polish with different graininess of abrasive are provided.

Keywords: polishing, corundum ceramics, diamond suspension, fabrics for polish.

ИССЛЕДОВАНИЕ ТЕХНОЛОГИИ ФОРМИРОВАНИЯ РЕЛЬЕФА РАДИОПОГЛОЩАЮЩИХ ПОЛИМЕРНЫХ КОМПОЗИТОВ

Филиал Федерального государственного унитарного предприятия «Российский федеральный ядерный центр – Всероссийский научно-исследовательский институт экспериментальной физики» «Научно-исследовательский институт измерительных систем им. Ю.Е. Седакова», г. Нижний Новгород

В статье представлены результаты по исследованию технологии формирования рельефа пластин высокоэффективного радиопоглощающего полимерного материала с углеродными наноразмерными трубками для диапазона электромагнитного излучения частотой 30-37 ГГц. Найдена оптимальная форма поверхности покрытия и проведена оценка радиопоглощающих характеристик разработанного материала. Показана стабильность характеристик листов радиопоглощающего материала к климатическим факторам.

Ключевые слова: поглощение электромагнитного излучения, радиопоглощение, радиопоглощающий материал, материал с углеродными нанотрубками, коэффициент стоячей волны по напряжению, формирование рельефа.

В связи с активным развитием радиоэлектронной аппаратуры и увеличением ее мощности всё острее встает вопрос обеспечения электромагнитной совместимости внутренних элементов приемопередающих устройств без увеличения их массы и габаритов. Как правило, достаточной мерой для решения проблемы электромагнитной совместимости является приклеивание тонких листов из радиопоглощающих материалов (РПМ) на внутренние поверхности ВЧ-блоков, что способствует устранению переотражения излучения от металлических стенок корпуса. Изготовление таких листов в большинстве случаев осуществляется на основе полимерной матрицы с электропроводными наполнителями [1]. В рамках проводимой работы к разрабатываемому материалу предъявлялись следующие требования:

- диапазон рабочих температур от -60 до +125°C;
- РПМ должен изготавливаться в виде листовых заготовок толщиной до 2 мм;
- коэффициент стоячей волны по напряжению (КСВн) листа РПМ на металлической подложке должен быть не более 2 в полосе частот шириной не менее 2 ГГц из диапазона 30 – 37 ГГц;
- КСВн листа РПМ должен оставаться в указанных пределах после воздействия климатических факторов, имитирующих 18-летнее складское и 3-летнее полевое хранение.

В качестве матрицы для получения материала с заданными свойствами был выбран силиконовый каучук СКТН с теплостойкостью до 200°C, а в качестве наполнителя, обеспечивающего радиопоглощающие свойства, - углеродные нанотрубки «Таунит-М» и полые стеклянные микросферы [2]. С целью улучшения технологических свойств, а также прочностных и радиопоглощающих характеристик разрабатываемого материала выбранные наполнители подвергались сушке в течение 2 часов при 110°C. На следующем этапе работы проводилось компьютерное моделирование РПМ с целью выбора оптимальной формы поверхности. Для достижения минимального КСВн при создании РПМ часто прибегают к формированию на поверхности РПМ рельефа, способствующего уменьшению отражения. Принцип такого подхода – увеличение количества переотражений падающей нормально к поверхности электромагнитной волны перед её отражением обратно в свободное пространство, представленного на рис. 1, на котором приведен наиболее классический пример формирования на поверхности РПМ рельефа в виде пирамид.

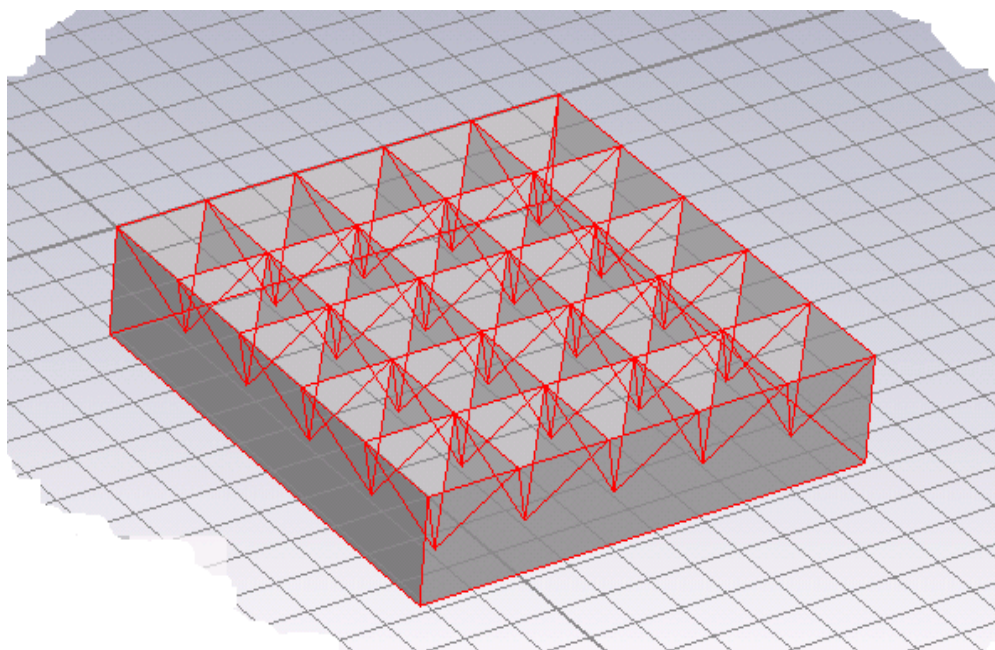


Рис. 1. 3D-модель РПМ с пирамидальными углублениями

С технологической точки зрения гораздо легче изготовить оснастку для формирования на поверхности РПМ углублений в виде пирамид, при том что эффект от создания рельефа (пирамид из РПМ или пирамидальных углублений) будет одинаков. С целью определения оптимальной формы и размеров пирамидальных углублений в РПМ использовали компьютерное моделирование в среде CST Microwave Studio. Моделировался участок волновода, нормально к оси которого размещался лист РПМ толщиной 2 мм с металлической подложкой. Поскольку определить значения диэлектрической проницаемости (ϵ) и тангенса угла диэлектрических потерь ($\text{tg}\delta$) исследуемых материалов не представлялось возможным из-за отсутствия соответствующего оборудования, для моделирования использовали оценочные значения данных величин: $\epsilon=4,5$; $\text{tg}\delta=0,35$. Схематично лист РПМ в разрезе приведен на рис. 2.

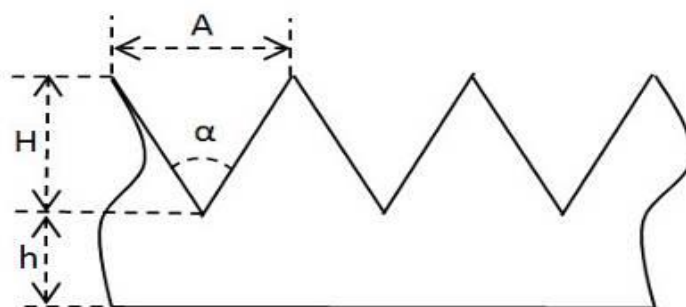


Рис. 2. Лист РПМ в разрезе

В результате моделирования рассчитывалась зависимость КСВн РПМ от частоты в диапазоне 30 – 37 ГГц. Для оценки использовали максимальное значение КСВн в данном диапазоне. Для поиска оптимального рельефа листов РПМ рассчитывали значение КСВн, изменяя размер основания (A) и угол (α) в вершине пирамиды. Отметим, что задавались такие значения α и A , при которых толщина основания листа РПМ под вершиной пирамиды (h) была не меньше 0,2 мм, т.е. высота пирамиды (H) не превышает 1,8 мм. Полученные данные приведены в таблице 1.

Таблица 1

Зависимость максимального КСВн в диапазоне 30 – 37 ГГц от размера и формы пирамидальных углублений в листе РПМ

Угол (α), °	Сторона основания пирамиды (А), мм	Высота пирамиды (Н), мм	КСВн
44	1,25	1,55	2,95
48	1,25	1,40	3,15
53	1,25	1,25	3,4
60	1,25	1,08	3,65
70	1,25	0,89	4
90	1,25	0,63	4,4
44	1,43	1,77	2,7
48	1,43	1,61	2,9
53	1,43	1,43	3,1
60	1,43	1,24	3,4
70	1,43	1,02	3,8
90	1,43	0,72	4,25
53	1,67	1,67	2,8
60	1,67	1,45	3,1
70	1,67	1,19	3,5
90	1,67	0,84	4,1
60	2	1,73	2,75
70	2	1,43	3,15
90	2	1,00	3,85
70	2,5	1,79	2,75
90	2,5	1,25	3,5
80	3	1,79	2,75

Для формования листов из РПМ была спроектирована и изготовлена пресс-форма, предназначенная для формования пластин размером 120x120 мм из РПМ методом литьевого прессования. Опираясь на данные таблицы 1, были выбраны два варианта геометрии пластин из РПМ с наименьшим значением КСВн и изготовлены соответствующие вкладыши в пресс-форму:

1. сторона основания пирамиды 2,00 мм, угол в вершине пирамиды 60°;
2. сторона основания пирамиды 1,43 мм, угол в вершине пирамиды 44°.

После проведения серии экспериментов по исследованию радиопоглощающих свойств РПМ различных составов было установлено, что лучшие свойства показывают образцы с углублениями в виде пирамид со стороной основания 2,0 мм и углом в вершине 60°. Также была измерена фактическая толщина получаемых листовых заготовок. Толщину листовых заготовок из полученного РПМ измеряли по 9-ти точкам. Результаты измерений приведены в таблице 2.

Толщина листовой заготовки составляет $1,54 \pm 0,18$ мм ($d_{cp} \pm 3S$), т.е. менее 2 мм. Величину разброса по толщине $\pm 0,18$ мм рекомендуется использовать также в качестве допуска на толщину листового материала при выпуске конструкторской документации.

Результаты измерений толщины РПМ

№ п/п	Толщина, мм	Среднее значение (d_{cp}), мм	Стандартное отклонение (S), мм
1	1,50	1,54	0,06
2	1,45		
3	1,59		
4	1,50		
5	1,50		
6	1,61		
7	1,60		
8	1,55		
9	1,60		

Для исследования стойкости РПМ к воздействию климатических факторов подвергли образцы материала следующим воздействиям:

- имитация 18-летнего складского хранения (для отапливаемых помещений);
- имитация 3-летнего хранения под навесом;
- воздействие максимальной рабочей температуры 125°C в течение 5 ч.
- после климатических испытаний вновь провели измерение КСВн РПМ на металлической подложке.

Все образцы радиопоглощающего материала после климатических испытаний имеют КСВн менее 2 в диапазоне частот 30,0 – 35,9 ГГц, что удовлетворяет предъявляемым требованиям. Таким образом, разработанный РПМ и покрытие на его основе (см. рис. 3) полностью удовлетворяют предъявляемым требованиям.



Рис. 3. Лист РПМ, отпрессованный в пресс-форме с вкладышем

Выводы

1. Разработан новый радиопоглощающий материал на основе силиконового каучука и углеродных нанотрубок.

2. Разработана форма радиопоглощающего покрытия, обеспечивающая достижение повышенных радиопоглощающих свойств.

3. Материал РПМ и покрытие на его основе полностью удовлетворяет предъявляемым требованиям, а именно:

- имеет диапазон рабочих температур от -60 до +125°C;
- из РПМ можно изготавливать листовые заготовки толщиной до 2 мм;
- РПМ в виде пластин с углублениями в форме пирамид со стороны основания 2,0 мм и углом в вершине 60, имеет КСВн на металлической подложке не более 2 в полосе частот шириной не менее 2 ГГц из диапазона 30 – 37 ГГц;

- КСВн данного РПМ остаётся в указанных пределах после воздействия климатических факторов, имитирующих 18-летнее складское, 3-летнее полевое хранение, воздействие максимальной рабочей температуры +150°C в течение 5 ч.

4. Разработана технологическая инструкция «Радиопоглощающий материал РПМ. Приготовление, прессование деталей и заготовок», что позволило расширить возможности технологической базы [3] «НИИИС им. Ю.Е. Седакова» для технологического обеспечения проектирования и изготовления СВЧ устройств [4].

Библиографический список

1. Михайлин Ю.А. Специальные полимерные композиционные материалы / Ю.А. Михайлин - СПб.: Научные основы и технологии. - 2008.
2. Бакина Л.И. Радиопоглощающий материал с низким уровнем отражения / Л.И.Бакина, В.Л.Зефирова, Е.А.Захарычев // Антенны. – 2016. - №1. -С. 45-48.
3. Светлаков Ю.А. Формирование технологической базы при технологическом обеспечении проектирования и изготовления СВЧ устройств / Ю.А. Светлаков // Проектирование и технология электронных средств. – 2018. - № 4. - С. 30-36.
4. Козлов В.А. Развитие структуры технологического обеспечения при системном проектировании и изготовлении СВЧ компонентов бортовых радиоэлектронных систем / В.А. Козлов, Ю.А. Светлаков, А.Ю. Седаков // Проектирование и технология электронных средств. - 2018. - № 3. - С. 23-29.

L.I. Bakina

RADAR-ABSORBENT POLYMER COMPOSITES PATTERNING TECHNOLOGY ANALYSIS

Branch of the Federal State Unitary Enterprise “Russian Federal Nuclear Center - All-Russian Scientific Research Institute of Experimental Physics” “Scientific Research Institute of Measuring Systems named after Yu.Ye. Sedakov”, Nizhny Novgorod

This paper provides research results for patterning of the high-efficiency radar-absorbent polymer plates integrated with carbon nanoscale tubes designed for 30-37 GHz band electromagnetic interference, identifies optimum-scale pattern of the coating surface and describes radar-absorbent properties of the developed material, including behavioral robustness against various climatic factors.

Keywords: electromagnetic absorption, radar absorption, radar-absorbent material, carbon nanotubes-integrated material, voltage standing wave ratio, patterning.

ИССЛЕДОВАНИЕ ТЕХНОЛОГИИ ФОРМИРОВАНИЯ LTCC-СТРУКТУР ДЛЯ СОЗДАНИЯ НЕМАГНИТНОГО КОРПУСА

Филиал Федерального государственного унитарного предприятия «Российский федеральный ядерный центр – Всероссийский научно-исследовательский институт экспериментальной физики» «Научно-исследовательский институт измерительных систем им. Ю.Е. Седакова», г. Нижний Новгород

В статье обсуждается вопрос возможности изготовления многослойной керамической платы толщиной 5 мм с полостью глубиной 2,5 мм без деформации, являющейся основанием для немагнитного корпуса.

Ключевые слова: многослойная керамическая плата, технология низкотемпературного совместного обжига, технология LTCC, коэффициент усадки, LTCC-основание, температурный профиль, немагнитный корпус.

В настоящей работе решалась задача изготовления немагнитного корпуса для создания трехосевого датчика магнитного поля. Для этого была выбрана технология низкотемпературного совместного обжига (Low Temperature Cofire Ceramic – LTCC), основными достоинствами которой являются [1]:

- проектирование и производство 3-D контуров;
- возможность размещения пассивных компонентов внутри подложки, что уменьшает размер контуров более чем на 50%, по сравнению с печатными платами;
- каждый слой контролируется до сборки модуля и, при необходимости, может быть заменен, что повышает процент выхода годных изделий;
- технологический цикл многослойных керамических модулей экологически чист и компактен;
- сокращение производственных циклов по сравнению с обычными толсто пленочными технологиями;
- многие процессы могут быть автоматизированы при серийном производстве;
- температура обжига порядка 850 °С позволяет применять материалы с малым удельным сопротивлением, такие как золото и серебро (для LTCC), вместо молибдена и вольфрама, которые используются в высокотемпературной технологии (HTCC);
- возможность вырезания ленты/подложки любой формы;
- применение немагнитных материалов позволяют использовать технологию в решении задач по магнитометрии.

Благодаря перечисленным достоинствам LTCC технология успешно применяется производителями многослойных плат для ВЧ и СВЧ корпусов микросхем и электронных приборов. LTCC материалы имеют неоспоримый потенциал развития и совершенствования для того, чтобы составить серьезную конкуренцию многослойным платам из стеклотекстолита и высокотемпературной керамики.

В НИИИС им. Ю.Е. Седакова (далее НИИИС) имеется опыт проектирования LTCC-устройств, создан участок по изготовлению МКП, оснащенный современным технологическим оборудованием. Маршрут изготовления МКП включает в себя следующие технологические операции: пробивка переходных отверстий и полостей, заполнение переходных отверстий пастой, трафаретная печать проводниковых элементов, пакетирование, ламинирование керамических заготовок, резка и обжиг модулей [2].

До настоящего времени на предприятии была разработана технология изготовления МКП толщиной до 3 мм (14 слоев) с неплоскостностью в пределах 50 мкм. В рамках реализации проекта по созданию немагнитного корпуса для размещения

магниторезистивного чувствительного элемента необходимо было освоить изготовление LTCC-основания, являющимся по сути МКП толщиной до 5 мм с полостью глубиной до 2,5 мм. Для проведения исследований был выбран керамический материал SK-47.

Основными операциями, которые определяют возможность изготовления образцов такой толщины, являются операции ламинирования и обжига.

При проверке стандартного режима операции ламинирования заготовок, не имеющих полостей, толщиной 5 мм были обнаружены трещины на торце образца. Поэтому режим ламинирования был изменен. А именно, увеличены такие параметры как время прогрева керамического пакета и температура ламинирования, что позволило получить образцы без расслоений и трещин.

Для проведения операции обжига был проведен анализ температурного профиля, рекомендованного специалистами фирмы КЕКО для структур толщиной более 5 мм [1]. Профиль приведен на рис. 1

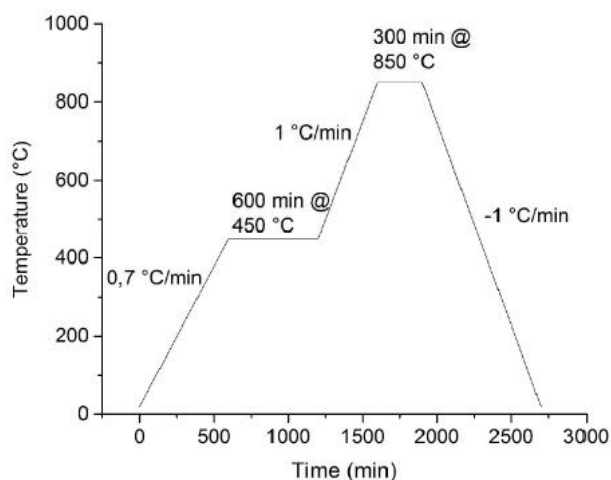


Рис. 1. Температурный профиль для структур толщиной более 5 мм

Из рис.1 видно, что длительность данного обжига—более 40 часов (около 2 сут.). Организация данного процесса при соблюдении его непрерывности, вызывает определенные сложности. Поэтому, была проведена консультация с производителем материалов на предмет разбивки обжига на несколько дней, исходя из максимального рабочего времени – 12ч. Специалисты компании «АК Микротех» (официальный представитель фирмы КЕКО) поддержали данную идею. Был разработан новый профиль обжига с разбивкой по этапам. Обжиг образца МКП по разработанному профилю проводился за 3 рабочих смены.

Первый этап – выжигание связки при температуре 450 °C. Данный этап был разбит на 2 дня по 12 часов каждый (включая время подъема температуры, выдержки при максимальном значении температуры и остывания).

Второй этап – обжиг образца при температуре 850 °C.

С целью исследований технологии изготовления МКПс заданными характеристиками были собраны 2 образца с полостями из 28 слоев керамики, дно полости -14 слоев. Размер образцов 23,5×19,0 мм, полостей – 6,7×6,4 мм. Для изготовления данных образцов были использованы разработанные ранее программы пробивки полостей.

Операцию ламинирования образцов проводили по разработанному режиму. Визуальный осмотр показал, что образцы не имеют трещин, деформации. Полученная толщина необожженных образцов составила ~ 6,16 мм, глубина полости ~3,25 мм.

При проведении резки керамической заготовки была увеличена температура нагрева стола и ножа. При проведении контроля была отмечена ровность края МКП, вырывов, сколов, трещин не обнаружено.

Обжиг первого образца был проведен за три дня в соответствии с разработанным профилем обжига. Производитель материалов не рекомендует при проведении обжига использовать нагрузку. Исходя из ранее проведенных работ по изготовлению МКП из материала SK-47 была определена нагрузка $2,2 \times P$, где P – масса образца. Роль нагрузки выполняли фрагменты сеттера (алюмооксидная керамика). Необходимо отметить, что для нагрузки выбирали наиболее подходящий по весу имеющийся в наличии фрагмент сеттера. Масса груза составила 16,26 г.

Визуальный осмотр показал, что после обжига неплоскостность образца – менее 20 мкм, других деформаций и трещин не наблюдали. Толщина образца после обжига составила 5,01 мм, глубина полости 2,55 мм. Внешний вид 28-слойного образца после обжига показан на рис. 2.

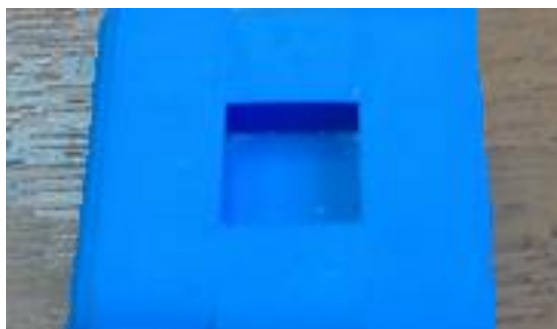


Рис. 2. Внешний вид 28-слойного образца после обжига

Значения коэффициента усадки и плотности приведены в таблице 1.

Таблица 1

Значения коэффициента усадки и плотности 28-слойного образца из керамики SK-47

	Коэффициент усадки K_{uc}		Плотность, ρ , г/см ³	
	По X, Y	По Z	Образец	По данным ТУ 6365-001-05717914-2019
Образец	По данным ТУ 6365-001-05717914-2019	Образец	По данным ТУ 6365-001-05717914-2019	Образец
1,141	1,156 (13,5±0,5) %	1,23	1,19 (16±2) %	2,935 не менее 2,9

Из таблицы 1 видно, что полученные значения коэффициента усадки отличаются от заявленных значений в ТУ [3]. Незначительное расхождение значений коэффициента усадки, полученного и заявленного в ТУ, объясняется различием используемого оборудования. Полученная плотность полностью соответствует требованиям ТУ.

Для подтверждения полученных результатов был проведен обжиг второго образца, аналогично первому. Визуальный осмотр показал, что образец после обжига также не имеет деформации, неплоскостность менее 20 мкм, трещины отсутствуют. Толщина образца после обжига составила 5,01 мм, глубина полости 2,55 мм. Полученный коэффициент усадки данного образца 1,143, немного отличается от коэффициента усадки предыдущего обжига 1,141, что можно объяснить погрешностью измерений.

Таким образом, проведенные исследования показали возможность изготовления образца МКП высотой 5 мм с полостью глубиной до 2,5 мм из керамики SK-47 фирмы КЕКО с использованием разработанных режимов ламинирования и обжига.

Для выполнения малогабаритного немагнитного LTCC-основания (МКП) была разработана конструкторская документация (КД). Конструктивно данная МКП размером $(20 \times 18) \pm 0,3$ мм состоит из 28 керамических слоев с полостью прямоугольной формы размером (15×11) мм и глубиной в 12 слоев, ширина стенки корпуса – 2,5 мм. Внутри большой полости расположены еще 3 небольшие прямоугольные полости глубиной 2 и 3

слоя. Отличительной особенностью данной МКП является торец со ступенью шириной 1 мм и высотой 14 слоев по длинной стороне образца.

На керамической заготовке размером 127×127 мм было размещено 12 модулей. На стыке 2-х модулей (по длинной стороне) образовалась полость размером (20×2) мм глубиной 14 слоев.

Были изготовлены фотошаблоны, программы пробивки отверстий и полостей с учетом коэффициента усадки 1,142 (среднее значение, полученное из предварительных экспериментов).

Для проведения операции трафаретной печати был выполнен комплект сетчатых трафаретов в количестве 9 шт. с капиллярной пленкой – Capillex.

Операции пробивки отверстий и полостей, заполнения отверстий проводниковой пастой, трафаретной печати проводили стандартным образом [2].

Операцию ламинирования проводили по разработанному для 28-слойного образца режиму с использованием стандартных силиконовых пластин, применяемых для ламинирования МКП с полостями. Последующий визуальный осмотр показал, что образцы с обратной стороны деформации не имеют, с лицевой стороны дно полостей также без деформации, но была выявлена деформация стенок по длинной стороне МКП, они выгнулись наружу. Внешний вид керамической заготовки после ламинирования приведен на рис.3.

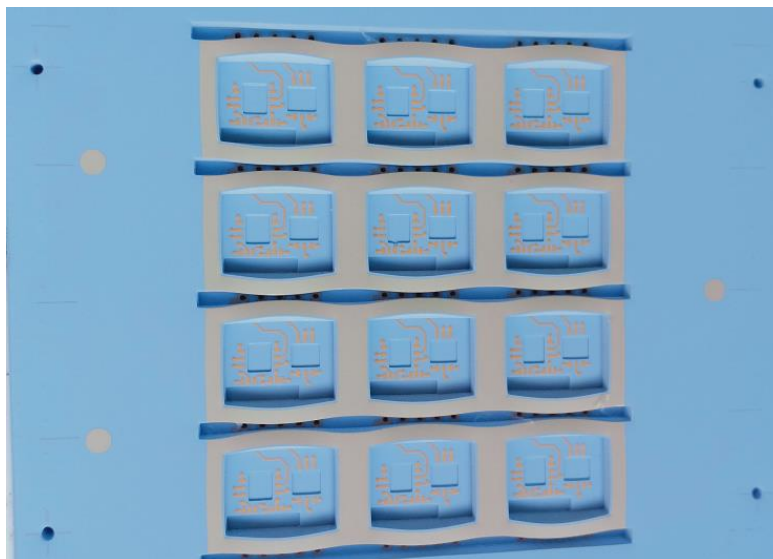


Рис. 3. Внешний вид керамической заготовки после ламинирования

Резку на модули проводили по опробованным режимам для 28-слойного образца.

Обжиг проводили по программе, разработанной для 28-слойного образца, так как это было описано ранее. Общее время обжига составило 34 ч. Данную операцию проводили также под нагрузкой $2,2 \times P$, где P – масса образца. С лицевой и обратной стороны образцы защищали «жертвенным» материалом.

Визуальный осмотр показал, что операция обжига не изменила внешний вид образцов, деформация стенки полости не увеличилась. Полученные значения коэффициента усадки 1,147-1,148 незначительно отличаются от заявленного значения в ТЗ 1,142 на УП для изготовления фотошаблонов, пробивки отверстий и полостей. Полученные габаритные размеры соответствуют требованиям КД. Неплоскостность образцов составила менее 50 мкм.

Внешний вид изготовленных МКП приведен на рис. 4.

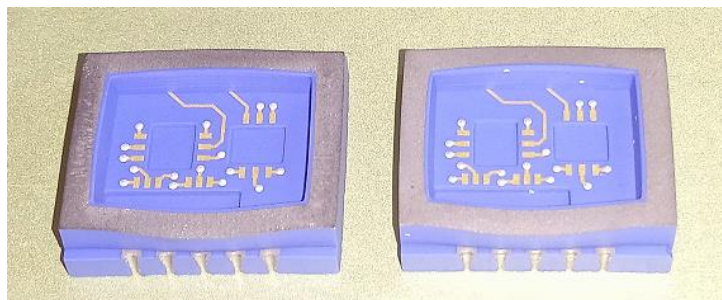


Рис. 4. Внешний вид LTCC-основания

Электронный контроль показал наличие всех электрических цепей.

В ходе проведенной работы был разработан маршрут изготовления и определены режимы ламинирования, резки, обжига для изготовления LTCC-основания толщиной 5 мм. Полученные образцы соответствуют требованиям КД по габаритным размерам и имеют неплоскостность менее 50 мкм. Деформации дна полостей, расслоения основания не обнаружено. Отмечена деформация 2-х стенок полости по длинной стороне корпуса. Решение данной проблемы требует дополнительной проработки, а именно, поиска технологических решений формирования узкой и глубокой полости, а также разработки технологических ограничений для ширины стенки полостей в зависимости от размеров полости. Таким образом, получена принципиальная возможность изготовления LTCC-основания толщиной 5 мм с полостью глубиной 2,5 мм для последующего размещения магниторезистивных чувствительных элементов для создания трехосевых датчиков магнитного поля.

Библиографический список

1. Правила проектирования многослойных LTCC/HTCC компонентов (модулей) // ООО «АК Микротех» (ред. 08.2019).
2. Гладышева, М.П. Исследования и разработка технологии изготовления многослойных керамических LTCC-плат из отечественной системы материалов / М.П. Гладышева, Е.В. Колпакова, Д.Е. Орехов, // Сборник материалов XXV Международной научно-технической конференции «Информационные системы и технологии – 2019» – Нижний Новгород: НГТУ им. Р.Е. Алексеева, 2019. - С. 173-176
3. Технические условия ТУ 6365-001-05717914-2019. Ленты LTCC/KEKOSK-47 // ООО «КЕКО-Р». - 2019.

М.П. Gladysheva, E.V. Kolpakova, D.E. Orehov, V.E. Sergeev

RESEARCH OF THE TECHNOLOGY OF FORMING LTCC STRUCTURES FOR CREATING A NON-MAGNETIC HOUSING

Branch of the Federal State Unitary Enterprise “Russian Federal Nuclear Center - All-Russian Scientific Research Institute of Experimental Physics” “Scientific Research Institute of Measuring Systems named after Yu.Ye. Sedakova”, Nizhny Novgorod

The article discusses the possibility of manufacturing a multilayer ceramic board with a thickness of 5 mm with a cavity depth of 2,5 mm without deformation, which is the basis for a non-magnetic housing.

Keywords: multilayer ceramic board, Low Temperature Co-firing Technology, LTCC technology, shrinkage coefficient, LTCC base, temperature profile, non-magnetic housing.

ТЕХНОЛОГИЯ ПОЛУЧЕНИЯ ТВЕРДОГО ПАЯЕМОГО ГАЛЬВАНИЧЕСКОГО ПОКРЫТИЯ СТОЙКОГО К МОРСКОМУ ТУМАНУ

Филиал Федерального государственного унитарного предприятия «Российский федеральный ядерный центр – Всероссийский научно-исследовательский институт экспериментальной физики» «Научно-исследовательский институт измерительных систем им. Ю.Е. Седакова», г. Нижний Новгород

Разработан технологический процесс нанесения на стальную и алюминиевую основы покрытия сплавом олово-никель, обладающего высокой твердостью, износостойкостью, хорошей паяемостью и высокой коррозионной стойкостью. Определены оптимальный состав электролита и режимы осаждения сплава. Установлены устойчивость покрытия к воздействию климатических факторов и стойкость к соляному (морскому) туману.

Ключевые слова: сплав олово-никель, хлористо-фтористый электролит, защита от коррозии, морской климат, СВЧ техника.

Гальванические покрытия в ряде случаев должны обеспечивать надежную защиту элементов радио электронной аппаратуры (РЭА) в жестких коррозионных условиях, эксплуатируемых в условиях морского климата [1]. Для защиты в данных условиях ранее применялось кадмиевое гальваническое покрытие, получаемое в цианистом электролите. Однако, в виду высокой токсичности цианистых электролитов, технологические процессы нанесения покрытий с экологической точки зрения не перспективны. Помимо кадмиевого покрытия также широко используются и цинковые покрытия. Поскольку коррозионная стойкость и защитная способность цинковых покрытий значительно ниже, чем кадмиевых, актуальной является проблема повышения коррозионной стойкости цинковых покрытий за счет их легирования другими металлами, и в первую очередь металлами подгруппы железа (никелем, кобальтом, железом). В результате такого легирования модификация свойств цинковых покрытий может быть достигнута как за счет формирования более тонкой совершенной структуры осадка с малым количеством легирующего металла, так и в результате осаждения более коррозионностойких фазовых структур–сплавов Zn-Ni, Zn-Co, Zn-Fe.

Возникла необходимость в поиске альтернативного защитного покрытия с возможностью его применения в новой разрабатываемой СВЧ-технике, с целью увеличения эксплуатационных характеристик в условиях высокой коррозионной нагрузки (таких как высокие или низкие температуры, повышенная влажность, соляной туман и др.). Альтернативным вариантом, указанных выше покрытий, являются электрохимические покрытия сплавом олово-никель, с возможностью эксплуатации в морском тумане, обладающее высокой стойкостью к внешним воздействующим факторам (ВВФ), пластичностью и паяемостью. Данное покрытие возможно получить по основным базовым материалам: углеродистые стали и алюминиевые сплавы.

По составу сплав олово-никель представляет собой интерметаллическое соединение. Стандартные электродные потенциалы олова и никеля различаются примерно на 0,1 В, причем потенциал никеля отрицательнее потенциала олова. Катодная поляризация по никелю выражена резче, чем по олову (при выделении их из простых солей) и, следовательно, потенциал никеля при электролизе становится еще более отрицательным. Поэтому совместное выделение олова и никеля возможно, очевидно, только при наличии таких комплексообразователей, которые резко снижали бы активность ионов олова и тем самым повышали потенциал их разряда. Из всех

комплексообразователей, которые были исследованы, наиболее подходящими оказались фториды, образующие комплексные соединения олова.

В результате информационного поиска были выявлены хлористо-фтористые электролиты с содержанием никеля в покрытии 30-35 %. В целях проведения сравнительных исследований было приготовлено 3 состава электролитов. Для приготовления растворов использовали реактивы марки «ч» и дистиллированную воду. Составы электролита олово-никель представлены в табл. 1.

Таблица 1

Электролиты для нанесения сплава олово-никель

Компоненты электролитов, г/л	№ электролита		
	1	2	3
Никель (II) хлорид 6-водный	250-300	250-300	250-300
Олово (II) хлористое 2-водное	45-50	40-50	40-50
Натрий фтористый	-	50-60	25-30
Аммоний фтористый	60-65	-	25-35

Для выбора оптимального состава электролита исследовали влияние на внешний вид катодного осадка комплексообразователей - фторидов натрия и аммония, кислотности электролита, температуры и плотности тока.

В результате исследований определен оптимальный состав хлористо-фтористого электролита:

- никель (II) хлорид 6-водный 250-300 г/л;
- олово (II) хлористое 2-водное 40-50 г/л;
- натрий фтористый 25-30 г/л;
- аммоний фтористый 25-35 г/л;

и режимы осаждения покрытия сплавом олово-никель:

- температура электролита 50-55 °С;
- плотность тока 0,5-2,5 А/дм²;
- значение рН 4,3-4,8.

Для исследования микрораспределения сплава использовали стандартную ячейку Хулла (см. рис. 1) с рабочим объемом 1000 мл в качестве источника тока применяли выпрямитель типа Б5-46.

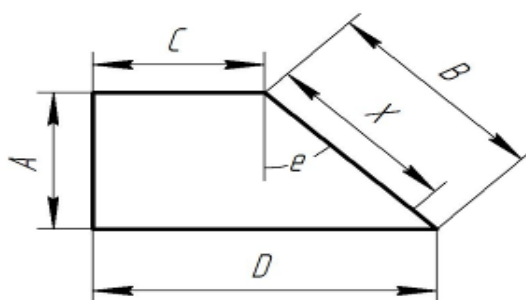


Рис. 1. Схема электролитической ячейки Хулла: А – анод; В – катод; X - длина полученного покрытия хорошего качества на катоде; С, D – боковые стороны

Допустимую катодную плотность тока D_k , А/дм² в зависимости от заданной силы тока I , А, и расстояния – X , мм для ячейки Хулла $V=1000$ мл, определяли по формуле (1):

$$D_k = 5,1I [1 - 0,434 \ln(X/10)]. \quad (1)$$

По отработанным режимам осаждения были нанесены покрытия сплавом олово-никель с массовой долей олова 65% толщиной 9 мкм с подслоями никеля и меди толщиной 6 мкм каждого с обозначением Н6.М6.О-Н(65)9 и Хим.Н6.М6.О-Н(65)9 на экспериментальные образцы из стали и из алюминиевых сплавов для последующего исследования паяемости, твердости и коррозионной стойкости покрытий.

Детали завешивали под нагрузкой, применяя повышенную плотность тока в начальной стадии электролиза (толчок тока) в течение 30 с в 2 раза превышающее рабочее значение. Скорость осаждения покрытия сплавом олово-никель при $D_k=1,0$ А/дм² составила 0,3 мкм/мин.

После нанесения покрытий на экспериментальные образцы оценивали качество покрытия по внешнему виду в соответствии с ГОСТ 9.301[2]. На поверхности покрытий экспериментальных образцов отсутствовали браковочные признаки: цвет покрытия сплавом олово-никель светло-серый.

Исследования паяемости [3] покрытия сплавом олово-никель проводили методом определения коэффициента растекания дозированной заготовки припоев ПОСК 50-18 и ПОС 61 в присутствии низкотемпературного флюса ФДФс по формуле (2):

$$K_p = \frac{S_p}{S_0}, \quad (2)$$

где K_p – коэффициент растекаемости;

S_p – площадь, занятая припоем после его расплавления и растекания, мм²;

S_0 – площадь, занятая припоем в исходном состоянии до его расплавления, мм².

Усредненные значения коэффициентов растекаемости вычисляли по результатам не менее трех опытов.

Коэффициент растекаемости равный единице и выше свидетельствовал о хорошей паяемости покрытия.

Исследование микротвердости покрытий сплавом олово-никель определяли методом вдавливания алмазной пирамиды [4] на твердомере КВ 30 по методу Виккерса при нагрузке 0,05 кг (0,4903 Н) и времени выдержки 10 с. Эксперимент проводили после 240 часового старения покрытий при комнатной температуре. Измерения проводили минимум в 3 точках с последующим усреднением данных. Микротвердость по Виккерсу HV вычисляли по формуле (3):

$$HV = 0,189 \frac{F}{d^2}, \quad (3)$$

где d – среднее арифметическое длин обеих диагоналей отпечатка, мм;

F – приложенная нагрузка, Н.

Толщина покрытия сплавом олово-никель для определения твердости составляла не менее 30 мкм. Микротвердость покрытия сплавом олово-никель составила 750 кгс/мм². Для сравнения: микротвердость кадмиевого покрытия 50 кгс/мм², микротвердость покрытия олово-висмут 20 кгс/мм².

Исследование коррозионной стойкости покрытий к внешним воздействующим факторам проводили по программе испытаний на стойкость к воздействию климатических факторов, имитирующих 18-ти летнее складское хранение в отапливаемом складе и 3-х летнее хранение на открытой площадке под навесом и на стойкость к воздействию соляного (морского) тумана. Испытания проводились на деталях с покрытиями олово-никель для РЭА, работающих в жестких условиях эксплуатации по 5-й и 6-й группам эксплуатации покрытий в соответствии ГОСТ 9-903[5]. Испытание на воздействие соляного тумана проводилось в 5 % растворе хлористого натрия при температуре 35 °С и влажности 98% в течении 7 суток. Оценивали внешний вид и состояние поверхности покрытий по ГОСТ 9.311 [6]. Коррозионных поражений покрытия и основного металла не

обнаружено. Покрытия сплавом олово-никель толщиной 9 мкм с подслоями никеля и меди по 6 мкм сохранили свои высокие декоративные и защитные свойства (по 10 баллов).

По результатам проведенных исследований и экспериментальных работ разработаны и внедрены технологические инструкции на электрохимическое осаждения сплава олово-никель на детали из стали и алюминиевых сплавов. Технология включает в себя основные операции:

1. Подготовка поверхности;
2. Электрохимическое и химическое (для сложнопрофильных деталей) осаждение никеля;
3. Электрохимическое осаждение меди;
4. Электрохимическое осаждение сплава олово-никель.

Технология рекомендуется к применению для получения как функциональных электропроводных, паяемых покрытий с высокой поверхностной износостойкостью взамен покрытия олово-висмут, так и для защиты от воздействия соляного тумана взамен кадмиевого покрытия.

Выводы

Выявлены, опробованы и испытаны перспективное электрохимическое покрытие сплавом олово-никель обеспечивающие наилучшие защитные свойства при эксплуатации РЭА СВЧ техники в условиях морского климата. Установлено, что исследуемое электрохимическое покрытие характеризуется высокой стойкостью к воздействию внешних климатических факторов, в том числе стойкостью к соляному туману, а также обладают хорошей паяемостью и могут быть рекомендованы к широкому применению в новой разрабатываемой СВЧ технике, эксплуатируемой в условиях морского климата.

Библиографический список

1. Голубев А.Н. Исследование порошковых покрытий для защиты изделий радиоэлектронной аппаратуры морского базирования/ А.Н. Голубев, В.Л. Зефирова// Материалы XXVI Международной научно-технической конференции «Информационные системы и технологии», ИСТ-2020. – Нижний Новгород, НГТУ им. Р.Е. Алексеева. – 2020. – С.1435-1439
2. ГОСТ 9.301-86 Покрытия металлические и неметаллические неорганические. Общие требования – [Текст]. – Введ. 1987.07.01. – М.: Межгосударственный совет по стандартизации, метрологии и сертификации: Стандартинформ, 2010.
3. ОСТ 4ГО.033.200 Припой и флюсы для пайки. Марка, состав, свойства и область применения – [Текст]. – Введ. 1980.01.01.
4. ГОСТ 9450-76 Измерение микротвердости вдавливанием алмазных наконечников. Требования – [Текст]. – Введ. 1977.01.01. – М.: Государственный комитет стандартов Совета Министров СССР: Издательство стандартов, 1976.
5. ГОСТ 9.303-84 Покрытия металлические и неметаллические неорганические. Общие требования к выбору – [Текст]. – Введ. 1985.01.01. – М.: Государственный комитет СССР по стандартам: Издательство стандартов, 1984.
6. ГОСТ 9.311-87 Покрытия металлические и неметаллические неорганические. Методы оценки коррозионных поражений – [Текст]. – Введ. 1988.07.01. – М.: Государственный комитет СССР по стандартам: Издательство стандартов, 1987.

A.N. Golubev, V. L. Zefirov, T.A. Zhirnova

**TECHNOLOGY FOR OBTAINING A SOLID SOLDERED ELECTROPLATING
COATING RESISTANT TO SEA FOG**

Branch of the Federal State Unitary Enterprise “Russian Federal Nuclear Center - All-Russian Scientific Research Institute of Experimental Physics” “Scientific Research Institute of Measuring Systems named after Yu.Ye. Sedakov”, Nizhny Novgorod

The technological process of applying a tin-nickel alloy coating on steel and aluminum bases, which has high hardness, wear resistance, good solderability and high corrosion resistance, has been developed. The optimal composition of the electrolyte and the deposition modes of the alloy are determined. The resistance of the coating to the effects of climatic factors and resistance to salt (marine climate) fog were established.

Keywords: tin-nickel alloy, chloride-fluoride electrolyte, corrosion protection, marine climate, superhigh-frequency devices.

РАЗРАБОТКА КОНСТРУКЦИИ И ТЕХНОЛОГИИ ИЗГОТОВЛЕНИЯ МАЛОГАБАРИТНЫХ МАГНИТОРЕЗИСТИВНЫХ ПРЕОБРАЗОВАТЕЛЕЙ

Филиал Федерального государственного унитарного предприятия «Российский федеральный ядерный центр – Всероссийский научно-исследовательский институт экспериментальной физики» «Научно-исследовательский институт измерительных систем им. Ю.Е. Седакова», г. Нижний Новгород

В статье описана разработка конструкции и технологии изготовления малогабаритного магниторезистивного преобразователя (МРП) на основе анизотропного магниторезистивного эффекта с характеристиками на уровне зарубежного аналога.

Ключевые слова: анизотропный магниторезистивный эффект, магниторезистивный преобразователь.

В последние десятилетия интерес к магниторезистивным преобразователям (МРП) существенно возрос благодаря таким свойствам и достоинствам, как широкий температурный диапазон, радиационная стойкость, возможность использования планарной интегральной технологии, совместимой с полупроводниковой технологией.

Прежде всего, для разработчиков измерительных систем представляют интерес первичные преобразователи магнитного поля на основе анизотропного и гигантского магниторезистивных эффектов. МРП на основе гигантского магниторезистивного эффекта обладают очень высокой чувствительностью к магнитному полю (десятки и сотни процентов), однако при этом требует высокого уровня технологической подготовки. В этих условиях разработка МРП на основе анизотропного магниторезистивного эффекта является наиболее перспективной. Анизотропные магниторезистивные чувствительные элементы имеют широкий рабочий диапазон, малый уровень шумов и небольшую стоимость производства. Чувствительные элементы на основе анизотропного магниторезистивного эффекта широко используются в системах определения (контроля) места положения и ориентации объекта, автоматического управления объектами различного назначения. Учитывая возросшие технические требования к современным системам управления, особенно в области навигации, требования по разрешающей способности МРП как к первичному преобразователю существенно возросли. Разрешающая способность определяется такими основными характеристиками МРП как удельная чувствительность, линейность передаточной характеристики и стабильность выходного напряжения.

Разрешающая способность МРП зависит как от конструкции (топологии) функциональных слоев, входящих в структуру, так и от технологии их изготовления. Поэтому, разработка конструкции и технологии изготовления МРП остается актуальным и в настоящее время.

В данной работе представлены результаты разработки конструкции и технологии изготовления МРП с габаритными размерами кристалла до $3,0 \times 3,0$ мм.

В основе принципа действия МРП лежит анизотропный магниторезистивный эффект, который заключается в способности длинной пермаллоевой (NiFeCo) пленки изменять свое сопротивление в зависимости от взаимной ориентации протекающего через нее тока и направления ее вектора намагниченности (рис.1) [1].

В качестве первичного преобразователя магнитного поля используется магниторезистивный мост Уитстона, состоящий из четырех идентичных групп магниторезистивных полосок, соединенных в меандр. Магниторезистивный мост Уитстона имеет ось легкого намагничивания, которая принудительно формируется в

процессе напыления магниторезистивной структуры во внешнем постоянном магнитном поле. МРП наиболее чувствителен к полям, направленным перпендикулярно этой оси.

Для получения линейной (или нечетной) передаточной характеристики МРП необходимо смещение вольт - эрстедной характеристики в линейную область. Наиболее предпочтительным способом смещения вольт – эрстедной характеристики МРП в линейную область является вариант с использованием полюсов Барбера (ПБ). Преимуществом данного способа является: меньшие геометрические размеры моста Уитстона и высокая стабильность выходного сигнала за счет формирования магниторезистивных полосок параллельно оси легкого намагничивания. Однако, данный способ имеет и свои технологические трудности в частности обеспечения надежного контакта (малого переходного сопротивления) между магниторезистивным участком и поводящим ПБ.

Для работы МРП в режиме максимальной чувствительности используется дополнительная пленочная катушка индуктивности «set/reset» (в русском переводе «установки/переустановки» или «сброса/установки»). Подача короткого установочного импульса тока 2 - 5 А длительностью 1 - 2 мкс через катушку индуктивности «set/reset» формирует поле, ориентирующее магнитные домены моста Уитстона в направлении легкой оси. Эта процедура позволяет перевести мост Уитстона в режим максимальной чувствительности и восстанавливает его характеристики, до воздействия на МРП сильным магнитным полем.

Катушка индуктивности «set/reset» выполняет еще одну важную функцию – это инвертирование передаточной характеристики МРП путем ее зеркального отображения. Инвертирование необходимо для реализации двухступенчатой методики измерения величины магнитного поля, которая исключает влияние температурного дрейфа элементов моста и схемы обработки сигнала, ошибку, вызванную нелинейностью характеристики, начальное смещение, а также потерю слабого сигнала на фоне сильных паразитных полей.

Катушка индуктивности «set/reset» выполняет еще одну важную функцию – это инвертирование передаточной характеристики МРП путем ее зеркального отображения. Инвертирование необходимо для реализации двухступенчатой методики измерения величины магнитного поля, которая исключает влияние температурного дрейфа элементов моста и схемы обработки сигнала, ошибку, вызванную нелинейностью характеристики, начальное смещение, а также потерю слабого сигнала на фоне сильных паразитных полей.

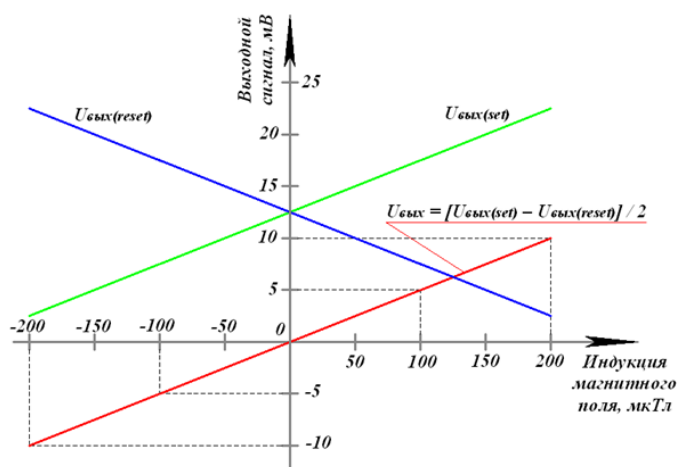


Рис. 1. Передаточная характеристика МРП

Методика включает три шага (рис. 1):

- формируется установочный импульс тока I_{set} , что обеспечивает условие «set».
- При этом измеряется и запоминается выходное напряжение моста $U_{вых}(set)$;
- формируется установочный импульс тока $I_{reset} = - I_{set}$, что инвертирует (зеркально отображает) характеристику преобразования. При этом измеряется и запоминается выходной сигнал $U_{вых}(reset)$;
- вычисляется по формуле:

$$U_{вых} = [U_{вых}(set) - U_{вых}(reset)] / 2 \quad (1)$$

Выбор конструкционных материалов

В качестве рабочей магниторезистивной структуры для создания моста Уитстона была выбрана структура FeNiCo – Ta – Al (где FeNiCo – сплав с химическим составом 65%Ni-15%Fe-20%Co). Напыление структуры производилось за один цикл откачки вакуумной камеры установки электронно-лучевого напыления «Оратория-9М».

В качестве межслойной изоляции между магниторезистивным мостом Уитстона и пленочной катушкой индуктивности «set/reset» использовался диоксид кремния SiO₂ толщиной (0,9 – 1,2) мкм формируемый плазмохимическим осаждением. Применение SiO₂ в качестве изоляционного слоя известно и широко представлено в литературе [2,3].

Разработка технологии изготовления МРП с уменьшенными размерами

Для разработки МРП с уменьшенными размерами, проводилась работа по проектированию тестовых структур МРП с различной шириной магниторезистивных полосок. В качестве аналога использовался МРП с шириной магниторезистивных полосок – 53 мкм и с габаритным размером 5,0x4,5 мм. В результате проектирования, были разработаны 5 вариантов исполнения с основными размерами представленными в таблице 1.

Таблица 1

Основные варианты тестовых структур МРП с уменьшенными топологическими размерами

№ испол.	Ширина магниторезистивной полоски, мкм	Зазор между магниторезистивными полосками, мкм	Ширина полюсов Барбера, мкм	Зазор между полюсами Барбера, мкм	Примечание
01	15	6,2	9	6	
02	18	7,5	10,8	7,2	
03	21	8,7	12,6	8,4	
04	24	10	14,4	9,6	
05	27	11,2	16,2	10,8	
06	53	30	24	20	Аналог

Изготовленные тестовые структуры МРП имели следующие параметры:

- сопротивление мостов Уитстона находится в диапазоне 900 – 1100 Ом;
- тестовые структуры МРП с шириной магниторезистивной полоски от 15 мкм до 27 мкм имеют нестабильность выходного сигнала от 2 до 4 мкВ, для структур с шириной магниторезистивной полоски 53 мкм характерны более высокие значения нестабильности выходного сигнала от 3 до 7 мкВ;
- значение удельной чувствительности тестовых структур изменяется как от ширины магниторезистивной полоски, так и от толщины магниторезистивного слоя (рисунок 2). С увеличением ширины магниторезистивной полоски и уменьшением толщины магниторезистивного слоя чувствительность увеличивается.
- габаритные размеры образцов составили 3,0x3,0 мм.

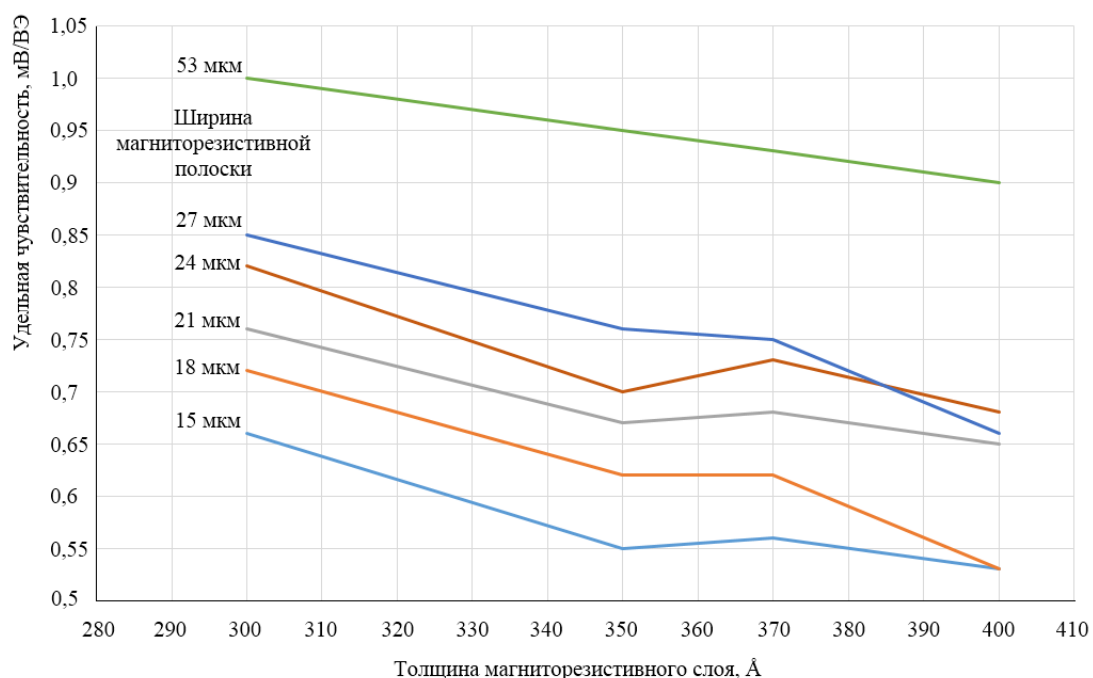


Рис. 2. Изменение удельной чувствительности тестовых структур от толщины магниторезистивной пленки и ширины магниторезистивных полосок

Из полученных данных видно, что наиболее перспективными вариантами дальнейшей проработки являются тестовые образцы с шириной магниторезистивных полосок 24 мкм и 27 мкм. Для повышения удельной чувствительности МРП дальнейшие исследования по разработке топологии проводились с учетом соотношения ширины ПБ к резистивному участку (расстоянию между ПБ) не менее 0,8, и с учетом минимального влияния размагничивающего фактора – величина зазора между магниторезистивными полосками должна быть минимальна: на рабочем шаблоне зазор должен составлять (2 – 4) мкм, а на изготовленных образцах (6 – 8) мкм.

На основе результатов исследований были разработаны образцы МРП с шириной магниторезистивных полосок 24 мкм и 27 мкм с учетом минимального влияния размагничивающих магнитных полей на выходные характеристики. Основные геометрические размеры элементов моста Уитстона представлены в таблице 2.

Таблица 2

Основные геометрические размеры моста Уитстона МРП.

Ширина магниторезистивной полоски, мкм	№ исполнения	Зазор между магниторезистивными полосками, мкм	Ширина полюсов Барбера (ПБ), мкм	Зазор между ПБ, мкм
24	-01	6,0	12,0	9,21
	-02	6,4	6,2	4,57
	-03	7,0	13	8,9
27	-04	7,0	14	10,0
	-05	7,0	7	5
	-06	8,0	15	9,75

С приведенными в таблице 2 значениями размеров элементов моста Уитстона, была изготовлена партия образцов МРП. Толщина магниторезистивного слоя варьировалась от 240 Å до 300 Å.

Результаты измерения МРП показали:

- значения сопротивлений мостов Уитстона находятся в диапазоне (990 – 1200) Ом, сопротивлений катушек индуктивности от 2 до 5 Ом;
- при увеличении толщины магниторезистивного слоя от 240Å до 300Å чувствительность МРП повышается (от ~ 0,84 мВ/ВЭ до 0,91 мВ/ВЭ для образцов с шириной магниторезистивной полоски 24 мкм, и от ~ 0,88 мВ/ВЭ до 0,97 мВ/ВЭ для образцов с шириной магниторезистивной полоски 27 мкм);
- при толщине магниторезистивной пленки (240 – 260) Å значение нестабильности выходного сигнала МРП составляет ~ (5 – 10) мкВ, а при толщине магниторезистивной пленки (270 – 300) Å нестабильность выходного сигнала составляет ~ (2 – 4) мкВ;
- наилучшие показатели были на образцах МРП исполнения №2 (ширина магниторезистивной полоски 24 мкм, чувствительность от ~ 0,87 мВ/ВЭ до 0,91 мВ/ВЭ, не более 4,5 мкВ) и №5 (ширина магниторезистивной полоски 27 мкм, чувствительность от ~ 0,90 мВ/ВЭ до 0,97 мВ/ВЭ, не более 4,0 мкВ).

Из полученных значений удельной чувствительности и нестабильности выходного сигнала можно сделать вывод, что наиболее оптимальная толщина напыления магниторезистивного слоя составляет (280 – 300) Å.

Перспективным вариантом изготовления являются образцы МРП исполнения №2 и №5. При высокой чувствительности они имеют небольшое значение нестабильности выходного сигнала (не более 4,5 мкВ).

Результаты измерения чувствительности и нестабильности выходного сигнала тестовых структур и МРП представлены в таблице 3.

Таблица 3

Результаты измерения чувствительности и нестабильности выходного сигнала тестовых структур и МРП

	Удельная чувствительность, мВ/ВЭ	Нестабильность выходного сигнала, мкВ
Тестовая структура 24 мкм	0,81 – 0,83	2,0 – 4,0
МРП 24 мкм	0,84 – 0,91	1,5 – 4,0
Тестовая структура 27 мкм	0,84 – 0,86	2,0 – 4,0
МРП 27 мкм	0,88 – 0,97	2,0 – 4,0

Из таблицы 3 видно, что МРП имеют более высокие значения удельной чувствительности по сравнению с тестовыми структурами. В среднем значение удельной чувствительности повышено на ~10 %. Значение нестабильности выходного сигнала находится на том же уровне по сравнению с тестовыми структурами.

Таким образом, в ходе исследований разработана конструкция и технология изготовления МРП с габаритными размерами 3,0×3,0 мм. Изготовленные МРП имеют высокую чувствительность (не менее 0,9 мВ/ВЭ) и небольшое значение нестабильности выходного сигнала (не более 4,5 мкВ) что соответствует зарубежному аналогу НМС 1021.

Библиографический список

1. Маргелов А. Магниторезистивные датчики положения компании Honeywell / А. Маргелов / Chipnews – инженерная микроэлектроника № 3, 2005, с60-65.
2. Patent US5952825 G01R 33/02, «Magnetic field sensing device having integral coils for producing magnetic fields», 14.08.1997.
3. Патент RU2066504 H01L 43/08, «Магниторезистивный датчик», от 20.07.1996.

S.V. Gorokhov, A.V. Negin, V.E. Sergeev

**WORKING OUT OF THE CONSTRUCTION AND MANUFACTURING TECHNIQUES
OF SMALL-SIZED MAGNETORESISTANCE TRANSDUCER**

*Branch of the federal universal enterprise «russian federal nuclear center – All – Russian
Scientific Research Institute of Experimental Physics» “Scientific Research Institute of
Measuring Systems named after Yu.Ye. Sedakov”, Nizhny Novgorod*

The article working describes the development of the design and manufacturing techniques of the small-sized magnetoresistance transducer based on an anisotropic magnetoresistive effect with characteristics at the level of a foreign analogue

Keywords: the anisotropic magnetoresistance, the magnetoresistance transducer.

Т.Л. Гудина, Ю.А. Светлаков

ИССЛЕДОВАНИЕ И ВЫБОР ТЕХНОЛОГИИ ФОРМИРОВАНИЯ ПОЛЯРИЗАЦИОННОЙ СТРУКТУРЫ ОТРАЖАТЕЛЯ ПАРАБОЛИЧЕСКОЙ ЗЕРКАЛЬНОЙ АНТЕННЫ

Филиал Федерального государственного унитарного предприятия «Российский федеральный ядерный центр – Всероссийский научно-исследовательский институт экспериментальной физики» «Научно-исследовательский институт измерительных систем им. Ю.Е. Седакова», г. Нижний Новгород

Представлены основные результаты исследований возможности получения поляризационной структуры отражателя параболической зеркальной антенны крайне высокочастотного диапазона радиоволн. Приводятся экспериментальные результаты изготовления поляризационной структуры антенны с применением электроэрозионной обработки и механической обработки резанием.

Ключевые слова: параболическая зеркальная антенна, поляризационная структура, электроэрозионная размерная обработка, фрезерование.

Рассматриваются технологии формирования поляризационной структуры отражателя параболических зеркальных антенн (ПЗА) крайне высокочастотного (КВЧ) диапазона радиоволн для бортовых радиолокационных систем, разрабатываемых с использованием технологической базы «НИИИС им. Ю.Е. Седакова».

Решаемая задача - разработать технологию, обеспечивающую получение изделия (в данном случае –ПЗА), не только функционально отвечающего своим требованиям (тактико-технические характеристики), но и требующую минимальных материальных и трудовых ресурсов на ее изготовление. Кроме того, согласно структуре технологического обеспечения проектирования и изготовления [1], проводилась работа по совершенствованию конструкции отражателя антенны на этапах технологической экспертизы при разработке конструкторской документации (КД), технологической подготовки производства и изготовления изделия на предмет технологичности изделия в условиях доступной технологической базы [2] и ее состояния на момент начала проектирования антенны.

К началу проектирования при анализе возможностей имеющихся технологических средств [3] было выявлено два возможных варианта формирования поляризационной структуры:

1. Вариант 1 - применение размерной электроэрозионной обработки (ЭЭО);
2. Вариант 2 - применение механической обработки резанием (фрезерование).

Неоднозначность выбора объясняется тем, что на тот момент времени отсутствовал опыт изготовления на предприятии аналогичных структур с заданными параметрами (отсутствие аналогов в составе технологической базы [1]). Следует отметить, что при проведении работ применение конкретного технологического оборудования продиктовано наличием его и других применяемых в технологиях технологических средств в составе технологической базы.

В рамках одной из НИР, после проведения расчетов (синтез и моделирование) была разработана КД на макет параболической зеркальной антенны, у которой поляризационная структура отражателя конструктивно представляет собой 189 канавок шириной 0,7 мм с шагом 1 мм и глубиной 2 мм на отражателе параболической формы поверхности диаметром 190 мм из алюминиевого сплава АМгб.

Технические требования, предъявляемые к поляризационной структуре согласно КД:

- материал отражателя – алюминиевый сплав АМг6;

- ширина канавок – $0,7^{+0,06}$ мм;

- шаг канавок - $1\pm 0,05$ мм;

- глубина канавок – 2 мм; теоретический профиль сечения канавок ограничен двумя параболоми, описываемых уравнениями $y^2=380x$ и $y^2=380(x+2)$ (поверхности Г и Е соответственно, см. рис. 1); среднеквадратичное отклонение поверхностей от теоретического контура должно составлять не более $\pm 0,15$ мм.

На рис.1 приведена конструкция отражателя ПЗА.

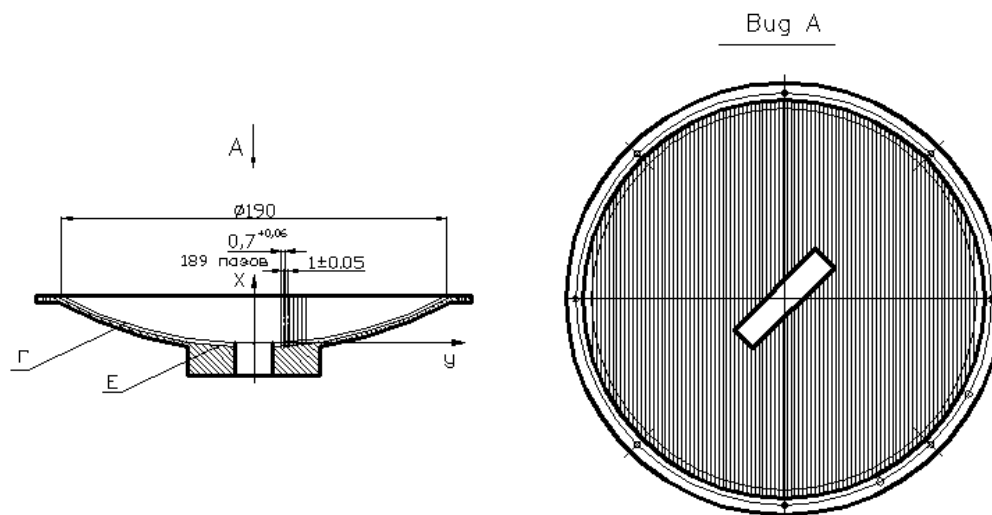


Рис. 1. Отражатель параболической зеркальной антенны

По первому варианту (метод ЭЭО) при изготовлении поляризационной структуры макета отражателя с диаметром поляризационной структуры 190 мм, обработка производилась на электроэрозионном координатно-прошивочном станке по управляющей программе (УП). Выбор параметров, автоматическое задание режимов обработки и генерация УП производилось в системе [4], входящей в состав программного обеспечения компьютерного числового программного управления (ЧПУ) станка, с моделированием процесса обработки непосредственно на станке [5]. Технологические режимы выбирались исходя из заданных технологических параметров (материала электрода, требуемой шероховатости, схемы обработки, глубины обработки, площади обработки).

Обработка производилась специально разработанным и изготовленным электродом-инструментом (ЭИ) из меди М1. Схема обработки объемным копированием на координатно-прошивочном станке показана на рис. 2.

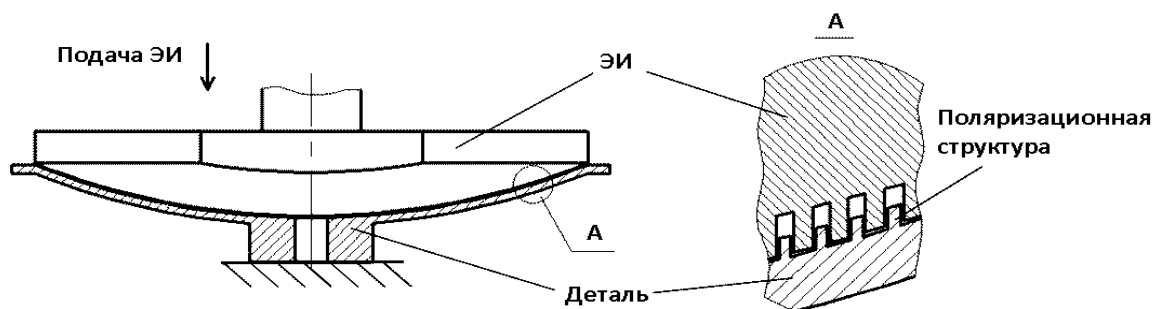


Рис. 2. Технологическая схема формообразования поляризационной структуры отражателя антенны с применением ЭЭО

Рабочую часть ЭИ, имеющую форму параболоида вращения (профиль сечения - парабола $y^2=380(x+2)$), изготавливали на токарном станке с ЧПУ. Формирование выступов на поверхности рабочей части ЭИ, предназначенных для объемного эквидистантного копирования на детали (для обработки канавок поляризационной структуры на координатно-прошивочном станке), производилась на вырезном электроэрозионном станке проволочным ЭИ диаметром 0,25 мм по разработанной для этого УП. Выбор параметров, автоматическое задание режимов обработки и генерация УП производились в системе «Автотехнолог», входящей в программное обеспечение компьютерного ЧПУ данного станка, исходя из задаваемых технологических параметров (материала заготовки, материала электрода-инструмента – проволоки, диаметра проволоки, высоты обработки, требуемой шероховатости обработки) [6]. Обработка выступов на параболоидной поверхности рабочей части ЭИ производилась последовательно с поворотом заготовки на станке с помощью универсального станочного приспособления.

Формирование структуры с применением метода ЭЭО де-факто обеспечивало получение данной структуры с заданными техническими требованиями, но это сопровождалось большими затратами на изготовление.

В процессе исследования возможности создания технологии с применением ЭЭО поляризационной структуры отражателя был выявлен ряд недостатков:

- высокая стоимость изготовления ЭИ, включая стоимость на основной материал (медь М1) расходных материалов и трудоемкость изготовления ЭИ;
- высокая стоимость ЭЭО поляризационной структуры (включая стоимость расходных материалов и трудоемкость ЭЭО);
- влияние износа ЭИ в процессе обработки на точность обработки поляризационной структуры (низкая стойкость инструмента – ограниченное число обрабатываемых изделий одним инструментом).

Это послужило основанием для поиска и выбора дополнительных технологических средств и технологий [3] для решения поставленной задачи. Был предложен второй вариант – формирование структуры механической обработкой резанием на станках с ЧПУ.

По второму варианту (механическая обработка резанием) при предварительном анализе КД и технологических возможностей метода фрезерования предложена технологическая схема формообразования канавок с помощью дисковой фрезы, которая должна иметь в процессе обработки траекторию подачи, повторяющую параболу согласно КД. Однако, в результате анализа, были выявлены следующие проблемы и недостатки формообразования канавок фрезерованием:

- высокая вероятность деформирования (отгиб) стенок канавок (из-за малых размеров - менее 0,3 мм) под действием поперечных сил резания в процессе обработки;
- относительно сложная УП для обработки всей поляризационной структуры отражателя;
- отклонение формы дна канавок от теоретического профиля.

Первые две из указанных проблем потребовали проведения работ по выбору размеров инструмента, режимов обработки, разработки УП с обеспечением заданной точности траектории движения инструмента в процессе обработки и др. с целью обеспечения точности изготовления поляризационной структуры согласно КД.

Третий недостаток (отклонение формы дна канавок от теоретического профиля) вызван тем, что форма канавок повторяет профиль инструмента. Так как по КД дно каждой канавки имеет индивидуальный профиль (что обеспечивает формообразование поверхности *E* согласно КД), то в идеальном случае необходимо проводить обработку каждой канавки фрезой с индивидуальной заточкой, повторяющей необходимый профиль дна канавки. Так, для обработки канавок отражателя необходимо иметь 95 различных фрез (или производить 95 высокоточных профильных перезаточек одной фрезы со

снятием и установкой инструмента на станке). Очевидно, что такой подход не перспективен, так как является очень затратным, трудоемким, имеет большой цикл обработки, и мала вероятность обеспечения требований КД из-за погрешностей, вызванных большим количеством переустановок и выверок инструмента. Обработка всех канавок одним инструментом (дисковой фрезой с прямоугольным профилем) позволяет избежать указанных проблем, но поверхность, образованная дном канавок (поверхность E отражателя), будет заведомо иметь отклонения от теоретического профиля.

Оценим отклонение формы поверхности E отражателя от теоретического профиля при обработке всех канавок одной фрезой.

На рис.3 показано поперечное сечение обрабатываемых канавок.

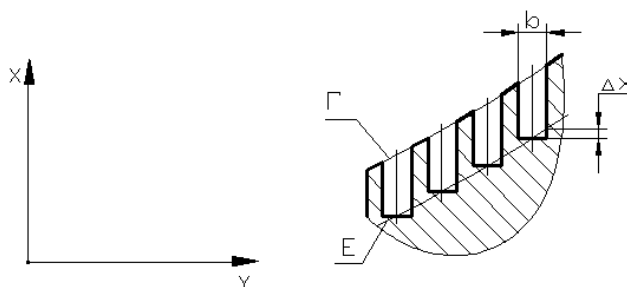


Рис. 3. Поперечное сечение канавок поляризационной структуры

Согласно КД, теоретический профиль сечения канавок ограничен поверхностями Γ и E . Считаем, что профиль канавки пересекается с теоретическим профилем дна канавок (поверхность E) по центру канавки, как показано на рис. 3. Тогда, зная ширину канавки b , можно определить величину максимального отклонения Δx для каждой i -той канавки из построений рисунка и уравнения для поверхности E :

$$\Delta x_i = \frac{(y_i + b/2)^2}{380} - \frac{y_i^2}{380}, \text{ или } \Delta x_i = \frac{by_i + (b/2)^2}{380},$$

где y_i – координата канавки по оси Y согласно КД; $i = 0, 1, 2, \dots, N$ (N – количество канавок).

Теоретически максимальное значение отклонения Δx составляет 0,173 мм (для 94-ой канавки). Среднее арифметическое значение $\overline{\Delta x}$ максимальных отклонений Δx_i от теоретического профиля всех канавок составляет 0,088 мм.

Так как оценка произведена для максимальных значений Δx каждой канавки, то очевидно, что при более строгой оценке (с учетом наличия, например, радиуса кромки инструмента, который к тому же увеличивается в процессе обработки в результате износа инструмента), ожидаемое среднее арифметическое значение величины Δx будет меньше. Данная оценка показывает, что при формировании поляризационной структуры будут выполнены (с большой вероятностью) требования КД (среднеквадратическое отклонение поверхностей от теоретического контура должно составлять не более $\pm 0,15$ мм).

Таким образом, оценка отклонения формы поверхности E отражателя от теоретического профиля показывает допустимость обработки всех канавок одной фрезой с прямоугольным профилем для обеспечения требований КД.

В результате экспериментальных исследований возможности изготовления поляризационной структуры методом фрезерования, в том числе оценки точности обработки, стойкости инструмента и подбора режимов обработки, удалось получить поляризационную структуру в соответствии с заданными техническими требованиями по КД.

Обработка поверхности Γ заготовки отражателя, имеющей форму параболоида вращения, производилась на токарном станке с ЧПУ.

Фрезерование канавок на параболической поверхности заготовки отражателя производилось на обрабатывающем центре по управляющей программе специально разработанной и изготовленной фрезой с использованием специальных оправок для придания жесткости инструмента и крепления его в шпинделе станка.

Основные режимы обработки, подобранные экспериментально:

- скорость вращения шпинделя 800 об/мин;
- скорость подачи режущего инструмента – 30 мм/мин;
- глубина резания – 2 мм (за один проход).

Все канавки были выполнены одной фрезой.

В процессе работы также были произведены измерения ширины и шага канавок, а также выборочные измерения глубины канавок на микроскопе УИМ-23. Контроль размеров показал, что размеры ширины и шага канавок соответствуют требованиям КД ($0,7^{+0,06}$ и $1\pm 0,05$ соответственно). Размеры глубины канавок составляют от 1,95 до 2,06 мм и соответствуют требованиям КД ($2\pm 0,15$).

Для определения точности обработки канавок поляризационной структуры была проведена статистическая обработка результатов измерений ширины и шага канавок. Результаты статистической обработки приведены в табл.1.

Таблица 1

Результаты статистической обработки размеров ширины и шага канавок отражателя

Статистические характеристики	Ширина канавки	Шаг канавок
Среднее арифметическое значение, \bar{X} , мм	0,7503	0,9991
Стандартное отклонение среднего значения, \bar{S} , мм	0,00066	0,00081
Доверительный интервал ($\bar{X} \pm 3\bar{S}$), мм	$0,7503 \pm 0,0020$	$0,9991 \pm 0,0024$
Требуемое по КД значение размера, мм	$0,7^{+0,06}$	$1 \pm 0,05$

Из данных, приведенных в таблице, видно, что фактическое среднее значение ширины канавок (0,7503 мм) смещено от середины поля допуска по КД (0,73 мм) на 0,0203 мм, фактическое среднее значение шага канавок (0,9991 мм) смещено от требуемого по КД (1 мм) на 0,0009 мм.

Анализ результатов, приведенных в табл. 1 показал, что доверительные интервалы ($\bar{X} \pm 3\bar{S}$) ширины и шага изготовленных канавок отражателя, находятся в пределах полей допусков размеров по КД.

В табл. 2 приведены сравнительные оценочные данные по затратам на инструмент (проектирование, основной материал, расходные материалы, изготовление и пр.) и изготовление поляризационной структуры (разработка технологии, разработка УП, трудоемкость обработки и т.д.) с применением методов ЭЭО и фрезерования (затраты для варианта с применением фрезерования принято за единицу).

Таблица 2

Сравнительные данные по затратам на изготовление поляризационной структуры

Показатели	ЭЭО	Фрезерование
Затраты на инструмент	~ 24	1
Затраты на обработку поляризационной структуры	~ 2,5	1
Общие затраты на формирование структуры	~ 8	1

Общие затраты (по оценочным данным) на изготовление поляризационной структуры отражателя методом фрезерования приблизительно в 8 раз ниже, чем методом ЭЭО.

Проведенные измерения электрических характеристик параболической зеркальной антенны с поворотом плоскости поляризации с отражателем, поляризационная структура которого изготовлена как методом фрезерования, так и с применением метода ЭЭО подтвердили их соответствие основным расчетным характеристикам.

В связи с положительными результатами измерений электрических характеристик и с учетом более низких материальных и трудовых затрат было рекомендовано применять метод фрезерования при изготовлении поляризационной структуры рассматриваемого отражателя, в результате чего были скорректированы требования КД (доработка КД на технологичность изготовления изделия) на отражатель в части формы канавок, формирующих поляризационную структуру.

Выводы

В результате проведенных исследований разработана технология, рекомендованная к использованию, и применяемая при изготовлении макетов отражателей параболических зеркальных антенн КВЧ диапазона радиоволн с параметрами поляризационной структуры, аналогичными рассмотренным, при проведении НИОКР.

Оценка затрат на формирование поляризационной структуры отражателя параболической зеркальной антенны показала целесообразность применения технологии механической обработки резанием (фрезерование на обрабатывающих центрах) канавок по сравнению с применением размерной электроэрозионной обработки.

Библиографический список

1. Козлов В.А. Развитие структуры технологического обеспечения при системном проектировании и изготовлении СВЧ компонентов бортовых радиоэлектронных систем / В.А. Козлов, Ю.А. Светлаков, А.Ю. Седаков // Проектирование и технология электронных средств. - 2018. - № 3. - С. 23-29.
2. Светлаков Ю.А. Формирование технологической базы при технологическом обеспечении проектирования и изготовления СВЧ устройств / Ю.А. Светлаков // Проектирование и технология электронных средств. – 2018. - № 4. - С. 30-36.
3. Светлаков Ю.А. Совершенствование и развитие технологической базы проектирования и изготовления СВЧ устройств / Ю.А. Светлаков // Материалы XXIII Международной научно-технической конференции «Информационные системы и технологии», ИСТ-2017. – Нижний Новгород, НГТУ им. Р.Е. Алексева. - 2017. - С. 1344-1349.
4. Светлаков Ю.А. Технология формирования апертуры волноводно-щелевой антенной решетки / Ю.А. Светлаков, А.Ю. Седаков // Проектирование и технология электронных средств. – 2019. - № 3. - С. 3-9.
5. Гудина Т.Л. Технология формирования апертуры волноводно-щелевой антенной решетки КВЧ диапазона / Т.Л. Гудина, Ю.А. Светлаков, А.Ю. Седаков // Труды XXIV Международной научно-технической конференции «Информационные системы и технологии», ИСТ-2018. – Нижний Новгород, НГТУ им. Р.Е. Алексева. – 2018. – С. 311-316.
6. Прокофьев Л.И. Методы изготовления волноводно-щелевых антенных решеток для бортовых АФС КВЧ диапазона длин волн / Л.И. Прокофьев, Ю.А. Светлаков, А.Ю. Седаков // Новые промышленные технологии. – 2009. – №3 – С. 11-14.

T.L. Gudina, Yu.A. Svetlakov

**RESEARCH AND CHOICE OF TECHNOLOGY OF FORMING POLARIZING
STRUCTURE OF THE REFLECTOR AT THE PARABOLIC ANTENNA**

Branch of the Federal State Unitary Enterprise "Russian Federal Nuclear Center - All-Russian Scientific Research Institute of Experimental Physics" "Scientific Research Institute of Measuring Systems named after Yu.Ye. Sedakov", Nizhny Novgorod

The main results of researches of a possibility of obtaining polarizing structure of the reflector of the parabolic antenna of extremely high-frequency range of radiowaves are provided. The experimental results of manufacture of polarizing structure of the antenna using electric erosion machining and milling are given.

Keywords: parabolic antenna, polarizing structure, dimensional electric erosion machining, milling.

РАЗРАБОТКА ТЕХНОЛОГИИ СТАБИЛИЗАЦИИ ПАРАМЕТРОВ ТОНКОПЛЕНОЧНЫХ РЕЗИСТИВНЫХ ЭЛЕМЕНТОВ

¹Филиал Федерального государственного унитарного предприятия «Российский федеральный ядерный центр – Всероссийский научно-исследовательский институт экспериментальной физики» «Научно-исследовательский институт измерительных систем им. Ю.Е. Седакова»,

²Нижегородский государственный технический университет им. Р.Е. Алексеева,
г. Н. Новгород

В работе разработана модифицированная технология, позволившая стабилизировать процессы внутри тонких резистивных пленок на основе сплава РС-3710, которые влияли на изменение проводимости, а также выявить скрытые дефекты в пленках на ранней стадии изготовления, и завершить диффузионные процессы, связанные с образованием окислов внутри пленок и диффузии активных металлов.

Ключевые слова: модификация, тонкопленочный резистивный элемент, стабилизация, сопротивление

Несмотря на достаточно высокий уровень развития технологии изготовления тонкопленочных резистивных элементов актуальной остается задача обеспечения стабильности их параметров.

Формирование наноразмерных пленок в вакууме происходит, как правило, за несколько секунд, и, следовательно, система «конденсат - подложка» оказывается существенно неравновесной [1, 2]. Исследование неравновесности системы «конденсат - подложка» раскрывает различные процессы, протекающие в тонких пленках под воздействием внешних факторов, которые изменяют параметры пленок [3]. Причины нестабильности заложены самой спецификой технологии тонкопленочной микроэлектроники, а элементы нестабильности возникают уже при эксплуатации [4].

Для стабилизации параметров тонкопленочных элементов используют методы модификации материалов, направленные на изменение их структуры и свойств.

Существует много способов модификации: термическая обработка в вакууме, инерционном газе, на воздухе, анодирование, электрическая нагрузка, лазерный отжиг. Суть этих методов сводится к уплотнению пленки и формированию равновесной структуры.

Одним из наиболее применяемым методом модификации является термическая обработка тонкопленочных структур, причем эффективнее воздействие термоциклических нагрузок. Механизмы термической стабилизации параметров сформированной тонкопленочной структуры определяются как минимум двумя основными процессами, протекающими внутри и на поверхности структуры:

- рекристаллизация микроструктуры тонкой пленки (рост зерен за счет их слияния в более крупные группы);
- отжига дефектов, срастания отдельных зерен при кристаллизации, окислении, фазовых переходах и уплотнения пленок.

Термическую обработку резисторов осуществляют двумя способами: отжигом в вакууме и на воздухе.

Отжиг в вакууме производят после напыления резистивной пленки; при этом за счет укрупнения кристаллитов происходит увеличение металлической проводимости, что вызывает смещение ТКС в сторону положительных значений и одновременное уменьшение ρ_s [5].

При термической обработке на воздухе повышение стабильности достигается за счет окисления пленки по границам зерен и по поверхности, что снижает интенсивность окислительно-восстановительных процессов, происходящих в пленке при ее эксплуатации. Окислительные процессы ведут к увеличению удельного сопротивления из-за образования поверхностного окисного слоя, образования изолирующих окисных межзеренных прослоек и увеличения относительной доли туннельного механизма проводимости по сравнению с металлической проводимостью [5].

В работе [6] путем проведения испытаний установлено, что стабильность пленок повышается при использовании комбинированного отжига вакуум-отжиг. Так как после отжига в вакууме пленка получается более плотной по сравнению с неотожженной, то окисление по границам зерен при отжиге на воздухе затруднено и преимущественно происходит рост поверхностной окисной пленки, толщина ее увеличивается, что служит хорошей защитой основного материала.

Операции термообработки в зависимости от решаемой задачи можно разделить на две основные группы: формирующий отжиг, в результате которого достигается требуемое сочетание величины удельного сопротивления (ρ_s) и температурного коэффициента сопротивления (ТКС), и стабилизирующий отжиг, обозначаемый также как искусственное старение, термотренировка, термостабилизация.

Условия отжига определяются рядом характеристик, к которым относятся, например:

1) давление (этот параметр может меняться от величин, характеризующих высокий вакуум, до значительно превышающих атмосферное – $10^{-3} - 2 \cdot 10^8$ Па);

2) наличие на поверхности обрабатываемой резистивной пленки дополнительных слоев, как изменяющих свои свойства в процессе отжига, так и лимитирующих воздействие внешней среды;

3) продолжительность отжига (как правило, основные изменения параметров пленки происходят в первые минуты отжига, а остальное время термообработки связано с релаксацией механических напряжений после завершения структурных превращений).

В настоящей работе исследованы резистивные тонкопленочные элементы из сплава РС-3710 на основе многокомпонентной системы хром-никель-кремний с удельным поверхностным сопротивлением ρ_s 1000 Ом/квadrat с контактами ванадий-алюминий, сформированные на подложках из ситалла марки СТ 50-1 методом ионно-плазменного распыления.

В процессе проведения исследований резистивные тонкопленочные элементы подвергались следующим видам технологических термотренировок:

- циклической криообработке при температуре 77 К в течение 1 мин и при 298 К в течение 1 мин, циклов не менее 3;

- отжигу на воздухе при температурах 573К, 673 К, 773 К в течение 1 часа;

- циклическому отжигу в вакууме. Циклический отжиг в вакууме проводился следующим образом: нагрев до 373К, затем в течение 30 мин рост температуры до 623К, и в течение (1,5 - 2,0) часов остывание до 373К. Количество циклов предусматривается не менее двух.

- комбинированному отжигу (циклический отжиг в вакууме + отжиг на воздухе).

По результатам исследований сопротивления резистивных пленок и его изменения построены кривые старения резистивных пленок, представленные на рис.1.

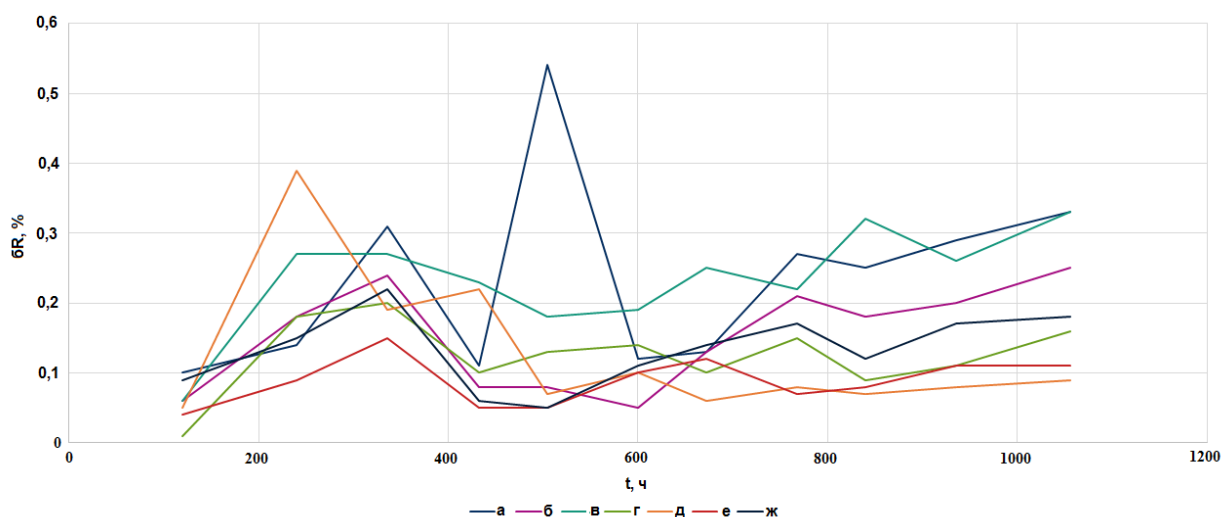


Рис. 1. Кривые старения резистивных пленок без обработки (а) и подвергнутых криообработке (б), отжигу на воздухе при $T = 573$ К (в), отжигу на воздухе при $T = 673$ К (г), отжигу на воздухе при $T = 773$ К (д), комбинированному отжигу вакуум – воздух при $T = 673$ К (е), комбинированному отжигу вакуум – воздух при $T = 773$ К (ж)

Анализ результатов показывает, что наилучшие результаты получены при комбинированном отжиге в вакууме и на воздухе при температуре 773 К. После циклического отжига в вакууме плёнка получается более плотной по сравнению с неотожжённой, следовательно, окисление по границам зёрен при отжиге на воздухе затруднено и одновременно с упорядочением структуры плёнки происходит рост поверхностной окисной плёнки.

В процессе отжига в пленке возникают внутренние напряжения, например, из-за несоответствия кристаллических решеток пленки и подложки, температурные напряжения из-за разности коэффициентов линейного расширения пленки и подложки, нестабильность структуры, все это влияет на старение пленок.

Таким образом, стабильность сопротивлений тонкопленочных резисторов подтверждает целесообразность применения комбинированного отжига с термоциклами в вакууме и отжигом на воздухе для завершения структурных превращений.

Библиографический список

1. Волохов И.В. Технологические методики повышения стабильности параметров тонкопленочных тензорезисторных датчиков давления: дисс. на соиск. уч. степ. канд. тех. наук/ И.В. Волохов.- М, 2008.
2. Аверин И. А. Влияние переходных процессов в тонкопленочной гетерогенной структуре на надежность чувствительных элементов тензорезисторных датчиков давления / И.В. Волохов, Е.А. Мокров, Р.М. Печерская // Известия высших учебных заведений. Поволжский регион. Технические науки. – 2008. – №2.– С. 123-127.
3. Гурин С.А. Методика термостабилизации тонкопленочных гетерогенных структур/ С.А. Гурин // Нано– и микросистемная техника. – 2014. – №11. – С. 12–16.
4. Миллер Ю.Г. Физические основы надежности интегральных схем./ Ю.Г.Миллер – М.: Сов.радио, 1976. – 318 с.
5. Ефимов И.Е. Микроэлектроника. Физические и технологические основы, надежность: Учеб. пособие для ВУЗов / И.Е. Ефимов, И.Я. Козырь, Ю.И. Горбунов– М.: Высшая школа, 1986. – 464 с.
6. Жебин А.П. Влияние термической обработки на электрические характеристики резистивных пленок/ А.П. Жебин// Электронная техника. - 1981. - выпуск 5 (108). - С. 28-33.

V.E. Sergeev^{1,2}

DEVELOPMENT OF STABILIZATION TECHNOLOGY PARAMETERS OF THIN-FILM RESISTIVE ELEMENTS

¹*Branch of the Federal State Unitary Enterprise «Russian Federal Nuclear Center – All-Russian Research Institute of Experimental Physics» «Yu. E. Sedakov Research Institute of Measurement Systems»*

²*Nizhny Novgorod state technical university named after R.E. Alekseeva, Nizhny Novgorod*

In this work, a modified technology was developed that allowed to stabilize the processes inside thin resistive films based on the RS-3710 alloy, which affected the change in conductivity, as well as to identify hidden defects in the films at an early stage of manufacture, and to complete the diffusion processes associated with the formation of oxides inside the films and the diffusion of active metals.

Keywords: modification, thin-film resistive element, stabilization, resistance

ПРИМЕНЕНИЕ КОМБИНИРОВАННОГО МЕТОДА ПОИСКА КОМПЛЕКСНЫХ КОРНЕЙ ДЛЯ РЕШЕНИЯ ДИСПЕРСИОННЫХ ЗАДАЧ ДЛЯ ОТКРЫТЫХ ЭЛЕКТРОДИНАМИЧЕСКИХ СТРУКТУР

¹Нижегородский государственный технический университет им. Р.Е. Алексеева

²Филиал Федерального государственного унитарного предприятия «Российский федеральный ядерный центр – Всероссийский научно-исследовательский институт экспериментальной физики» «Научно-исследовательский институт измерительных систем им. Ю.Е. Седакова»

Представлен комбинированный численный метод поиска комплексных корней, объединяющий в себе метод Мюллера и метод вариации фазы. Показаны результаты решений дисперсионных задач для цилиндрических направляющих структур, соответствующих различным типам волн, которые были получены с помощью данного метода.

Ключевые слова: численные методы, дисперсионные уравнения, диэлектрический волновод, комплексные волны, присоединенные волны, плазмон-поляритонные волны, нанопленки.

Введение

Исследование спектров волн неоднородных электродинамических направляющих структур сводится к решению краевых задач на однородном уравнении Гельмгольца относительно скалярных функций электромагнитного поля. Данные краевые задачи относятся к классу несамосопряженных краевых задач [1-3], собственные значения, которых в общем случае являются комплексными величинами.

При отсутствии диссипации энергии действительные собственные значения соответствуют поверхностным волнам, комплексные значения – вытекающим волнам, собственным и несобственным комплексным волнам. Однако возможность существования данных волн в конкретной структуре зависит от её параметров и рассматриваемого частотного диапазона. В связи с этим важными являются проблемы целенаправленного поиска комплексных решений дисперсионных уравнений, определения частотных диапазонов и областей параметров направляющей структуры, в которых существуют комплексные волны.

Методы поиска комплексных корней и оценка корректности найденных решений краевых задач

Поиск решений дисперсионных задач (в общем случае являющихся комплексными) для направляющих структур связан с использованием различных численных методов расчета, позволяющих получить приближенные решения уравнений с заданной точностью.

Одним из таких методов является метод Мюллера [4]. Метод Мюллера в отличие от других численных методов (секущих, хорд, Ньютона) использует не линейную, а квадратичную аппроксимацию функции. В нем для нахождения очередного приближения x_{k+1} необходимы три предыдущие точки: x_k, x_{k-1}, x_{k-2} .

В настоящее время метод Мюллера благодаря высокому быстродействию широко используется для поиска комплексных корней в различных САПР, предназначенных для расчета электродинамических структур. Однако, наряду с истинными нулями комплексной функции $F(\beta)$ при использовании этого метода может быть получено решение, соответствующее локальному минимуму, что приводит к появлению ложных корней.

Таким образом, метод Мюллера является приемлемым методом поиска комплексных корней в ограниченной области комплексной плоскости в небольшом частотном диапазоне, где отсутствуют локальные минимумы. В противном случае необходимо проверять истинность полученных решений другими методами, например, методом вариации фазы.

Метод вариации фазы [1, 5] позволяет получить истинные комплексные решения дисперсионного уравнения $F(\beta) = 0$ для широкого круга электродинамических задач, он основан на принципе аргумента:

$$\frac{1}{2\pi i} \oint_L \frac{F'(\beta)}{F(\beta)} d\beta = \frac{1}{2\pi} \text{var} [\arg F(\beta)]|_L. \quad (1)$$

Одновременное использование метода Мюллера и метода вариации фазы (комбинированный метод поиска комплексных корней) позволило использовать только лучшие стороны обоих методов, а именно: быстроту метода Мюллера и возможность идентификации истинных корней методом вариации фазы.

При решении электродинамических задач проекционными методами принципиальную роль в оценке адекватности полученных результатов играет исследование сходимости алгоритма: по волновым числам; по выполнению условий на границах частичных областей (при использовании метода частичных областей); по осуществлению предельного перехода к хорошо изученным структурам, а также по стремлению к нулю потока мощности собственных комплексных волн в среднем за период через поперечное сечение недиссипативной направляющей структуры [1, 6].

Вытекающие, собственные комплексные и медленные несобственные волны в круглом открытом диэлектрическом волноводе

С использованием комбинированного метода поиска комплексных корней, были получены решения дисперсионной задачи для круглого открытого диэлектрического волновода (КОДВ) [7]. На рис. 1а приведены дисперсионные характеристики и характеристики затухания симметричных и гибридных волн двухслойного КОДВ ($\epsilon_1 = 9,6$; $\epsilon_2 = 1,0$; $a_1 = 0,001$ м). По горизонтальной оси отложена нормированная частота $V = k_0 a_1 \sqrt{\epsilon_1 - \epsilon_2}$, где k_0 – волновое число плоской волны в свободном пространстве, a – радиус диэлектрического волновода; по вертикальной – нормированные действительная и мнимая части продольного волнового числа $\beta = \beta_1 + i\beta_2$.

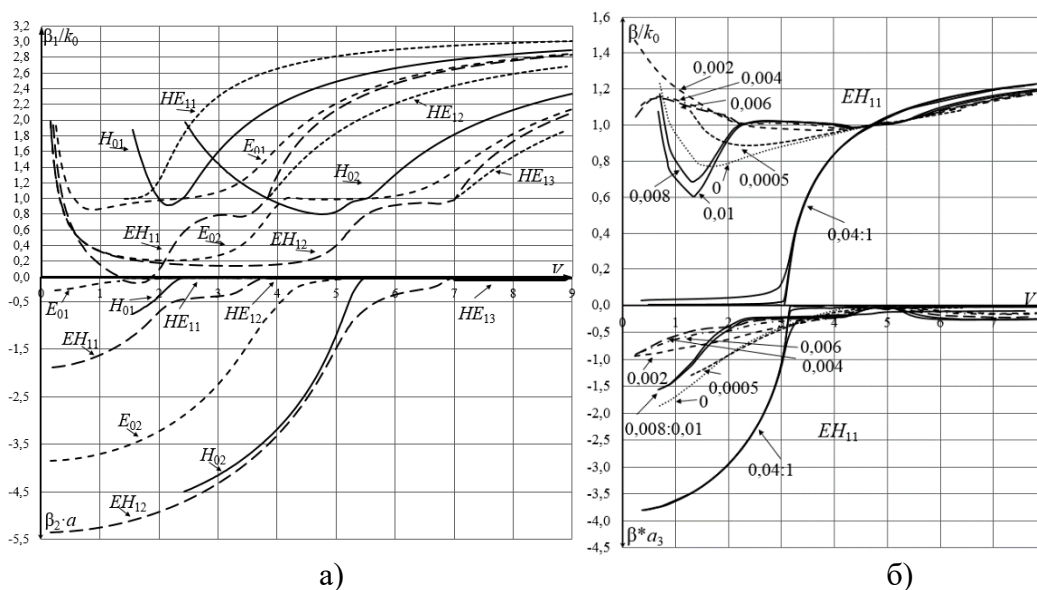


Рис. 1. Дисперсионные характеристики и характеристики затухания волн КОДВ без нанесения и с нанесением резистивной пленкой

На рис. 1а дисперсионные зависимости выше критической частоты (постоянная замедления волны $\beta_1/k_0 = 1$) соответствуют поверхностным волнам. На критических частотах поверхностные волны переходят в волны излучения – вытекающие волны. При переходе к вытекающим волнам поперечное волновое число находится в первом квадранте комплексной плоскости. Вытекающие волны при уменьшении частоты обычно переходят в медленные несобственные волны [1], у которых поперечное волновое число также находится в первом квадранте комплексной плоскости. На дисперсионных характеристиках волн EH_{nm} возможно существование участков, соответствующих собственным комплексным волнам. При этом поперечное волновое число переходит в четвертый квадрант комплексной плоскости. На рис. 1а такой участок существует у характеристики волны EH_{11} ($\beta_1/k_0 < 0$) в диапазоне частот $V \in [1,3 \div 1,8]$. При малых n это имеет место при достаточно больших отношениях диэлектрических проницаемостей слоев ϵ_1/ϵ_2 , в случае больших n – при меньших значениях ϵ_1/ϵ_2 .

У собственных комплексных и поверхностных волн напряжённость поля убывает при удалении от волновода, у медленных несобственных и вытекающих волн – нарастает. При покрытии КОДВ резистивной пленкой получены результаты, представленные на рис. 1б. На данном рисунке приведены дисперсионные характеристики и характеристики затухания волны EH_{11} четырехслойного КОДВ, покрытого резистивной пленкой ($\epsilon_1 = 2,1; \epsilon_2 = 1,3; \epsilon_3 = 1,79; \epsilon_4 = 1; a_1 = 4$ мкм, $a_2 = 4,5$ мкм, $a_3 = 62,5$ мкм). Числами обозначены значения поверхностной проводимости резистивной пленки. Для составления дисперсионного уравнения использовался метод поверхностного тока, в котором резистивная пленка представляется как токоведущая поверхность [7]. Для данного метода необходимо выполнение соотношения $\sigma \gg \epsilon \omega$.

Решения дисперсионного уравнения на комплексной плоскости поперечного волнового числа для поверхностных волн находятся в третьем квадранте в области медленных волн. С увеличением значения поверхностной проводимости резистивной пленки решения отдаляются от мнимой оси поперечного волнового числа. При уменьшении частоты поперечное волновое число стремится к точке начала координат и, проходя её, соответствует вытекающей волне, находящейся в первом квадранте в области быстрых волн. На низких частотах вытекающая волна EH_{11} сначала переходит в комплексную волну, а затем сначала в быструю, а далее с уменьшением частоты в медленную несобственную волну.

Присоединенные волны в круглом открытом диэлектрическом волноводе

Кроме дисперсионных задач, решения которых соответствуют нормальным волнам, для направляющих структур могут быть сформулированы присоединённые краевые задачи, состоящие из дифференциальных уравнений и систем граничных условий, которые образуются дифференцированием по собственному значению исходного дифференциального уравнения и системы граничных условий, соответственно. Решения присоединённых краевых задач соответствуют так называемым присоединённым волнам [1], характерной особенностью которых является наличие линейной зависимости их амплитуд от продольной координаты.

В работе [8] был сделан вывод о возможности существования присоединённых волн в КОДВ, описываемых решениями присоединённой краевой задачи на однородном уравнении Гельмгольца. Решение данного уравнения может быть найдено в виде:

$$\Pi_z^{e,m} = [R(r)f(z) + \bar{R}(r)\bar{f}(z)] \begin{cases} \cos n\varphi \\ \sin n\varphi \end{cases} . \quad (2)$$

Дисперсионная задача для присоединённых волн КОДВ образуется совокупностью трех трансцендентных уравнений, получаемых по методике, описанной в [8, 9].

На рис. 2 приведены решения данных трансцендентных уравнений: решения уравнения 1 – сплошная линия 1, соответствующая дисперсионной характеристике волны EH_{11} КОДВ; уравнения 2 – штриховая линия 2 и уравнения 3 – точечная линия 3. Точка А – общее решение трех уравнений. По горизонтали отложена нормированная частота $V = k_0 a \sqrt{\varepsilon_1 - \varepsilon_2}$, $\varepsilon_1 = 9$, $\varepsilon_2 = 1$. По вертикали отложена действительная часть нормированного продольного волнового числа $\tilde{\beta} = \beta / k_0$.

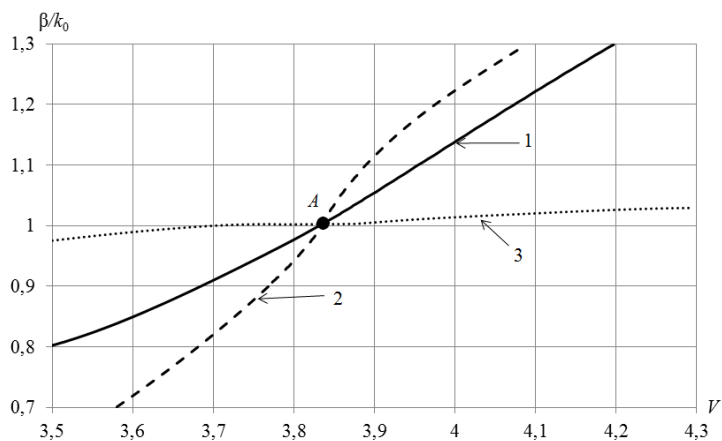


Рис. 2. Решения уравнений, образующих дисперсионную задачу для присоединённых волн

Численные исследования показали, что общее решение уравнений (точка А) существует и находится на нормированной частоте $V \approx 3,83$.

Результаты решения дисперсионной задачи в окрестности точки существования общего решения уравнений показали, что на частотах $V < 3,83$ значение поперечного волнового числа второй области α_2 находится в первом квадранте комплексной плоскости. Исследование особенностей поля в данной частотной области говорит о том, что напряженность поля при удалении от волновода возрастает. Следовательно, можно утверждать, что полученные решения соответствуют вытекающей волне.

Результаты решения дисперсионной задачи и расчеты распределений поля волны EH_{11} на частотах $V > 3,83$ показали, что волна EH_{11} в данной частотной области является поверхностной – напряженность поля экспоненциально убывает с удалением от волновода.

Таким образом, можно сделать вывод, что точка А, соответствующая решению дисперсионной задачи для присоединённых волн, является точкой перехода поверхностной волны EH_{11} в вытекающую волну. В данной частотной точке найденная присоединенная волна обеспечивает полноту системы волн КОДВ.

Плазмон-поляритонные волны в круглом открытом диэлектрическом волноводе с металлической нанопленкой

Хорошо известно, что металлы на оптических частотах обладают комплексной диэлектрической проницаемостью с отрицательной действительной частью и положительной мнимой частью. На границе раздела двух сред, диэлектрическая проницаемость, одной из которых является отрицательной, могут возникать поверхностные плазмон-поляритонные волны. Волны образуются при взаимодействии фотонов с электронным газом в металле. Особенностью этих волн является сильная локализация поля вблизи поверхности раздела.

При расчетах использовалась формула Друде, которая описывает изменение диэлектрической проницаемости металлической нанопленки от частоты.

На рис.3а изображены дисперсионные характеристики четной 1 и нечетной 2 поверхностных плазмон-поляритонных волн и комплексной волны 3 для структуры без учета потерь в пленке со следующими параметрами: радиус диэлектрического стержня $a = 100$ нм, толщина пленки 10 нм, $\epsilon_1 = 2,84, \epsilon_3 = 1$, материал пленки – серебро: $\epsilon_{r0} = 6, \epsilon_{i0} = 0$ [10]. На дисперсионной характеристике нечетной волны вблизи точки А имеется двузначный участок. В точке А образуется комплексная волна 3. Таким образом, при поиске решений дисперсионного уравнения для структуры без диссипации энергии (рис. 3а), также необходимо производить поиск волновых чисел на комплексной плоскости.

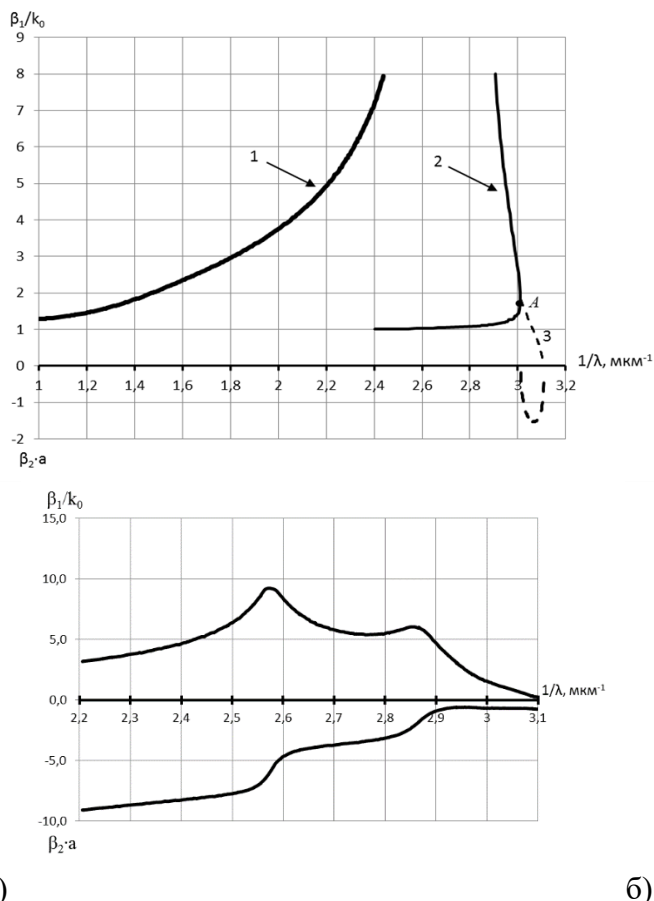


Рис. 3. Дисперсионные характеристики плазмон-поляритонных волн

На рис. 3б показаны дисперсионные характеристики и характеристики затухания для структуры со следующими параметрами: радиус диэлектрического стержня $a = 100$ нм, толщина пленки 10 нм, $\epsilon_1 = 2,84, \epsilon_3 = 1$, с учетом комплексности диэлектрической проницаемости металла.

Дисперсионная характеристика плазмон-поляритонной волны имеет два максимума, в случае, когда металлическая пленка окружена диэлектриками с разными диэлектрическими проницаемостями. Большой максимум (рис. 3б) расположен на частоте, на которой $|\epsilon_2| = |\epsilon_1| = 2,84$, а меньший максимум – на частоте, когда $|\epsilon_2| = |\epsilon_3| = 1$. Разрывов в характеристике нет, а участок, соответствующий комплексной волне в среде без потерь, является продолжением дисперсионной характеристики поверхностной плазмон-поляритонной волны на высоких частотах. Такая трансформация наблюдается и в других структурах с нанопленками [11]. Большой максимум находится на частоте $1/\lambda \approx 2,57 \text{ мкм}^{-1}$, к которой асимптотически приближается характеристика четной плазмон-поляритонной волны 1 (рис. 3а). Меньший максимум находится на частоте $1/\lambda \approx 2,85 \text{ мкм}^{-1}$, к которой асимптотически приближается характеристика нечетной плазмон-поляритонной волны 2 (рис. 3а).

Решения дисперсионного уравнения на комплексной плоскости поперечного волнового числа для комплексной волны находятся в третьем квадранте. При уменьшении частоты возникают поверхностные плазмон-поляритонные волны, у которых поперечное волновое число внешней области также находится в третьем квадранте. При дальнейшем уменьшении частоты поперечное волновое число стремится к точке начала координат и, проходя её, соответствует вытекающей волне, находящейся в первом квадранте.

Заключение

Большую роль при нахождении комплексных решений различных дисперсионных задач для направляющих структур играет выбор численного метода поиска решений.

Алгоритм расчета дисперсионных уравнений на основе комбинированного метода поиска комплексных корней, объединяющий в себе метод Мюллера и метод вариации фазы, показал свою эффективность и был применён для расчета дисперсионных характеристик различных волн цилиндрических направляющих структур. При этом обеспечивалась возможность определения положения поперечного волнового числа внешней области на комплексной плоскости, что необходимо для классификации типов волн.

Библиографический список

1. Раевский А.С. Комплексные волны / А.С. Раевский, С.Б. Раевский. – М.: Радиотехника, 2010. – 223 с.
2. Веселов Г.И. Слоистые металлodieлектрические волноводы / Г.И. Веселов, С.Б. Раевский. – М.: Радио и связь, 1988. – 248 с.
3. Малахов В.А. Комплексные волны в экранированной микрополосковой линии / В.А. Малахов, А.С. Раевский // Радиотехника и электроника. – 1999. – Т. 44. – № 1. – С. 58–61.
4. Muller D. E. A Method for Solving Algebraic Equations Using an Automatic Computer / D. E. Muller // Mathematical Tables and Other Aids to Computation. – 1956. – №10. – P. 208–215.
5. Бритов И.Е. Целенаправленный поиск комплексных волн в направляющих электродинамических структурах / И.Е. Бритов, А.С. Раевский, С.Б. Раевский // Антенны. – 2003. – В. 5(72). – С.64–71.
6. Малахов В.А. О решении дисперсионных уравнений волн направляющих электродинамических структур на комплексных плоскостях волновых чисел / В.А. Малахов, А.С. Раевский, С.Б. Раевский // Журнал вычислительной математики и математической физики. – 2015. – Т. 55. – № 6. – С. 1028–1038.
7. Малахов В.А. Расчет распределений плотностей потоков мощности волн в круглом ОДВ в закритической области / В.А. Малахов, А.А. Никитин, А.С. Раевский // Антенны. – 2017. – № 10. – С. 53–59.
8. Малахов В.А. Присоединённая волна круглого открытого диэлектрического волновода / В.А. Малахов, А.С. Нечаев, А.С. Раевский, С.Б. Раевский // Письма в журнал технической физики. – 2019. – Т. 45. – № 17. – С. 35–38.
9. Малахов В.А. Присоединённые волны в круглом двухслойном экранированном волноводе / В.А. Малахов, А.С. Раевский, С.Б. Раевский // Письма в журнал технической физики. – 2011. – Т. 37. – № 2. – С. 71–79.
10. Федянин Д.Ю. Поверхностные плазмон-поляритоны с отрицательной и нулевой групповыми скоростями, распространяющиеся по тонким металлическим пленкам / Д.Ю. Федянин, А.В. Арсенин, В.Г. Лейман [и др.] // Квантовая электроника. – 2009. – Т. 39. – № 8. – С. 745–750.
11. Малахов В.А. Комплексные волны и комплексный резонанс в структурах с металлическими нанопленками на оптических частотах / В.А. Малахов, К.В. Попков, А.С. Раевский // Труды НГТУ. – 2012. – №1. – С.120–127.

V.A. Malakhov¹, A.S. Nechaev^{1,2}, A.A. Nikitin², A.S. Raevskii¹

**APPLICATION OF THE COMBINED METHOD OF COMPLEX ROOTS FINDING
FOR DISPERSION PROBLEMS SOLVING FOR OPEN ELECTRODYNAMIC
STRUCTURES**

¹ *Nizhny Novgorod state technical university n.a. R.E. Alekseeva*

² *Branch of the Federal State Unitary Enterprise "Russian Federal Nuclear Center - All-Russian Scientific Research Institute of Experimental Physics" "Scientific Research Institute of Measuring Systems named after Yu.Ye. Sedakov", Nizhny Novgorod*

Purpose: Solving of various dispersion problems for a round open dielectric waveguide.

Design/methodology/approach: The combined numerical method of complex roots finding is presented. This method combines the Muller method and the phase variation method.

Findings: The results of dispersion problems solving for cylindrical guide structures corresponding to different types of waves obtained using this method are presented.

Research limitations/implications: To confirm the existence of complex waves in guide structures, the problems of finding the complex roots of dispersion equations are very important.

Originality/value: Use the combined method of complex roots finding allowed us to use only the best sides of the both methods.

Keywords: numerical methods, dispersion equations, dielectric waveguide, complex waves, adjoint waves, plasmon-polariton waves, nanofilms

ИССЛЕДОВАНИЕ ЯВЛЕНИЯ КОМПЛЕКСНОГО РЕЗОНАНСА С ИСПОЛЬЗОВАНИЕМ ЭНЕРГЕТИЧЕСКИХ ЛИНИЙ

Нижегородский государственный технический университет им. Р.Е. Алексеева

Природа собственных комплексных волн объясняется распределенным разворотом потока мощности к источнику, в результате чего возникает замкнутый поток мощности, который наблюдается при явлении комплексного резонанса, возникающего в результате взаимодействия двух комплексно-сопряженных комплексных волн. Предлагается использование энергетических линий для уточнения свойств комплексного резонанса.

Ключевые слова: комплексные волны, комплексный резонанс, энергетические линии.

В работах [1-4] показано, что в поперечно-неоднородных (слоистых) направляющих электродинамических структурах могут существовать, несмотря на отсутствие диссипации энергии, волны с комплексными волновыми числами – комплексные волны (КВ). В [1] отмечается, что источники, описываемые действительными функциями координат, возбуждают, в частности, в двухслойном круглом экранированном волноводе по обе стороны от себя по две комплексно-сопряженных волны с противоположно направленными фазовыми скоростями. Это приводит к возникновению стоячей волны, поле которой локализовано вблизи источника. При этом отрезок волновода, включаемый «на проход» или «на отражение» (в первом случае в плоскости симметрии, перпендикулярной оси волновода, располагаются возбуждающий и воспринимающий электроды, во втором – один возбуждающий электрод), во всем диапазоне существования комплексных волн ведет себя как резонатор и имеет при этом фильтрующие свойства. Поскольку отмеченное явление обнаруживает резонансные свойства (возрастание выходного сигнала в схеме «на проход» и резкое падение коэффициента стоячей волны в схеме «на отражение») не на одной частоте, как при обычном резонансе, а во всем диапазоне существования КВ, оно классифицировано [1, 5] как «комплексный резонанс» (КР).

Резонансным признаком рассматриваемого явления служит факт увеличения запасенной энергии в полосе частот, соответствующей диапазону существования комплексных волн, что позволяет ввести понятие добротности в энергетической формулировке и измерить ее. Экспериментальное исследование комплексного резонанса на примере круглого двухслойного экранированного волновода привели к созданию полосовых СВЧ фильтров простой конструкции с теоретически рассчитываемой полосой пропускания.

Исследование распределения электромагнитного поля в круглом двухслойном экранированном волноводе (КДЭВ) позволило обнаружить эффект резкого возрастания амплитуды поля при переходе из частотной области, соответствующей распространяющимся волнам, в область существования комплексных волн [1, 6-7]. Поскольку это явление по многим признакам аналогично резонансу в обычных резонансных структурах, в [5] это явление названо комплексным резонансом. КР отличается от обычного резонанса тем, что проявляется в широком диапазоне частот, соответствующем диапазону существования комплексных волн. Эта особенность обуславливает невозможность измерения добротности КР обычными резонансными методами. Добротность структуры, в которой возникает КР, удобнее оценивать путем сравнения ее с добротностью эталонного резонатора, работающего в том же диапазоне частот и дающего то же увеличение амплитуды поля при прохождении сигнала в схеме «на проход».

Широкополосность КР позволяет использовать его в датчиках, реагирующих на изменение поля в полосе частот и поэтому не требующих их постоянной подстройки. Пространственно-временные датчики на КР в силу их широкополосности могут реагировать на любую форму отраженного сигнала в радиометрическом канале радиоинтерферометра.

Рассмотрим особенности комплексного резонанса в отрезке двухслойного круглого экранированного волновода, нагруженного с двух сторон на законченные отрезки круглых однородно заполненных волноводов, рис. 1. Идеальным условием возбуждения комплексного резонанса, как отмечалось выше, является неограниченная протяженность направляющей структуры, однако реальные конструкции имеют конечную длину, и должны быть экранированы.

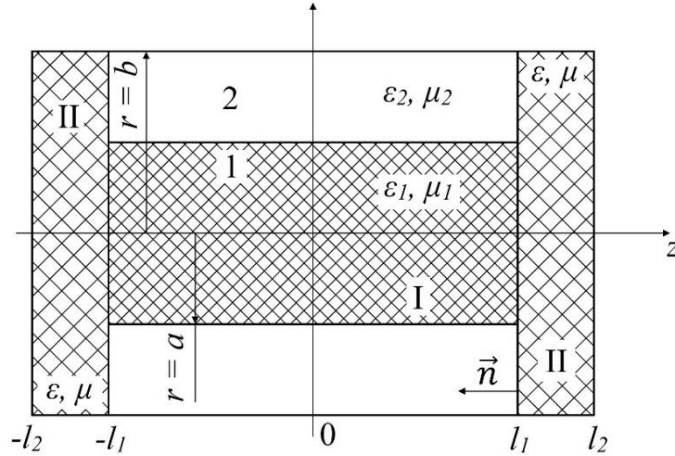


Рис. 1. Продольное сечение резонансной структуры

В силу симметрии системы рассматриваем ее половину, выделив справа от плоскости $z = 0$, в которой располагается источник, две области I – двухслойный волновод, II – закороченный отрезок однородно заполненного волновода.

Полагая, что источник возбуждает справа от себя две комплексных волны HE_{11} : прямую с амплитудой A и обратную с амплитудой $\bar{A} = -A^*$, продольные компоненты векторов Герца в областях I, II запишем в виде:

$$\begin{aligned}
 \Pi_{zI}^m &= A_{1,2} \psi_{1,2}^m(r) \cos \varphi e^{-i\beta z} + \bar{A}_{1,2} \bar{\psi}_{1,2}^m(r) \cos \varphi e^{-i\bar{\beta} z} + \\
 &R_{1,2} \psi_{1,2}^m(r) \cos \varphi e^{i\beta(z-l_1)} + \bar{R}_{1,2} \bar{\psi}_{1,2}^m(r) \cos \varphi e^{i\bar{\beta}(z-l_1)} + \\
 &+ \sum_k A_{1,2}^{(k)} \psi_{1,2}^{(k)m}(r) \cos \varphi e^{-i\beta_k z} + \sum_n R_{1,2}^{(n)} \psi_{1,2}^{(n)m}(r) \cos \varphi e^{-i\beta_n(z-l_1)}; \\
 \Pi_{zII}^e &= B_{1,2} \psi_{1,2}^e(r) \sin \varphi e^{-i\beta z} + \bar{B}_{1,2} \bar{\psi}_{1,2}^e(r) \sin \varphi e^{-i\bar{\beta} z} + \\
 &C_{1,2} \psi_{1,2}^e(r) \sin \varphi e^{i\beta(z-l_1)} + \bar{C}_{1,2} \bar{\psi}_{1,2}^e(r) \sin \varphi e^{i\bar{\beta}(z-l_1)} + \\
 &+ \sum_k B_{1,2}^{(k)} \psi_{1,2}^{(k)e}(r) \sin \varphi e^{-i\beta_k z} + \sum_n C_{1,2}^{(n)} \psi_{1,2}^{(n)e}(r) \sin \varphi e^{-i\beta_n(z-l_1)}; \\
 \Pi_{zII}^m &= \sum_p D_1^{(p)} J_1(\alpha_p^m r) \cos \varphi \sin \beta_p^m (z - l_2); \\
 \Pi_{zII}^e &= \sum_q D_2^{(q)} J_1(\alpha_q^e r) \sin \varphi \cos \beta_q^e (z - l_2),
 \end{aligned} \tag{1}$$

где $\beta = \beta_1 - i\beta_2$ – продольное волновое число прямой комплексной волны;
 $\bar{\beta} = -\beta_1 - i\beta_2$ – продольное волновое число обратной волны, $\beta_{1,2} \geq 0$;

$$\psi_{1,2}^e(r) = \begin{cases} J_1(\alpha_1 r), \\ \chi_{11}(\alpha_2 r), \end{cases} \quad \psi_{1,2}^m(r) = \begin{cases} J_1(\alpha_1 r), \\ \chi_{21}(\alpha_2 r), \end{cases}$$

$$\psi_{1,2}^{\varepsilon(k,n)}(r) = \begin{cases} J_1(\alpha_1^{k,n} r), \\ \chi_{11}(\alpha_1^{k,n} r), \end{cases} \quad \psi_{1,2}^{m(k,n)}(r) = \begin{cases} J_1(\alpha_2^{k,n} r), \\ \chi_{21}(\alpha_2^{k,n} r), \end{cases}$$

$$\chi_{11}(\alpha_2 r) = \frac{J_1(\alpha_2 r) Y_1(\alpha_2 b) - J_1(\alpha_2 b) Y_1(\alpha_2 r)}{J_1(\alpha_2 a) Y_1(\alpha_2 b) - J_1(\alpha_2 b) Y_1(\alpha_2 a)};$$

$$\chi_{21}(\alpha_2 r) = \frac{J_1(\alpha_2 r) Y_1'(\alpha_2 b) - J_1'(\alpha_2 b) Y_1(\alpha_2 r)}{J_1(\alpha_2 a) Y_1'(\alpha_2 b) - J_1'(\alpha_2 b) Y_1(\alpha_2 a)};$$

$J_1(\alpha_i r)$ и $Y_1(\alpha_i r)$ – цилиндрические функции 1-го и 2-го рода, соответственно,

$$\bar{\psi}^m(r) = \psi^{m*}(r); \bar{\psi}^\varepsilon(r) = -\psi^{\varepsilon*}(r);$$

$$\varepsilon_i \mu_i \omega^2 = \alpha_i^2 + \beta^2; \varepsilon_i \mu_i \omega^2 = (\alpha_i^*)^2 + \bar{\beta}^2;$$

$$\varepsilon_i \mu_i \omega^2 = (\alpha_i^{k,m})^2 + \beta_{k,n}^2; i = 1, 2;$$

$$\varepsilon \mu \omega^2 = (\alpha_{p,q}^{m,\varepsilon})^2 + (\beta_{p,q}^{m,\varepsilon})^2,$$

$\alpha_{1,2}$; $\alpha_{1,2}^{k,n}$ – корни дисперсионного уравнения соответственно комплексных и обычных волн двухслойного волновода;

α_p^m – p -й корень уравнения $J_1'(\alpha_p^m b) = 0$;

α_q^ε – q -й корень уравнения $J_1(\alpha_q^\varepsilon b) = 0$.

В записи (1) учли, что источник в общем случае наряду с комплексными волнами возбуждает обычные распространяющиеся и реактивно затухающие волны, а на стыке двухслойного и однородно заполненного волноводов так же возбуждаются как комплексные волны (прямые и обратные), так и обычные волны указанных волноводов.

$A_{1,2}$ и

$\bar{A}_{1,2}$ – амплитуды комплексных волн (прямой и обратной, соответственно), возбуждаемых источником; $R_{1,2}$ и $\bar{R}_{1,2}$ – амплитуды комплексных волн (прямой и обратной, соответственно), возбуждающихся на стыке волноводов; $A_{1,2}^{(k)}$ и $R_{1,2}^{(k)}$ – амплитуды обычных волн, возбуждаемых источником и отраженных от стыка; $D_{1,2}^{(p,q)}$ – амплитуды волн однородно заполненного волновода, возбуждающихся на стыке.

В соответствии с условием возбуждения

$$\bar{A}_{1,2} = -A_{1,2}^* \quad (2)$$

поля в области двухслойного волновода с учетом граничных условий при $r = a$ можно записать в виде

$$\begin{aligned} \vec{E}_I &= A_1 \vec{\xi}_0^{(-)} e^{-i\beta z} + \bar{A}_1 \vec{\xi}_0^{(-)} e^{-i\bar{\beta} z} + \sum_k A_1^{(k)} \vec{\xi}_k^{(-)} e^{-i\beta_k z} + \\ &+ R_1 \vec{\xi}_0^{(+)} e^{i\beta(z-l_1)} + \bar{R}_1 \vec{\xi}_0^{(+)} e^{i\bar{\beta}(z-l_1)} + \sum_n R_1^{(n)} \vec{\xi}_n^{(+)} e^{i\beta_n(z-l_1)}; \\ \vec{H}_I &= A_1 \vec{h}_0^{(-)} e^{-i\beta z} + \bar{A}_1 \vec{h}_0^{(-)} e^{-i\bar{\beta} z} + \sum_k A_1^{(k)} \vec{h}_k^{(-)} e^{-i\beta_k z} + \\ &+ R_1 \vec{h}_0^{(+)} e^{i\beta(z-l_1)} + \bar{R}_1 \vec{h}_0^{(+)} e^{i\bar{\beta}(z-l_1)} + \sum_n R_1^{(n)} \vec{h}_n^{(+)} e^{i\beta_n(z-l_1)}, \end{aligned} \quad (3)$$

где $\vec{\xi}_0^{(\pm)}$, $\vec{h}_0^{(\pm)}$ и $\vec{\xi}_0^{(\pm)}$, $\vec{h}_0^{(\pm)}$ – векторные функции полей прямой и обратной комплексных волн соответственно; $\vec{\xi}_k^{(-)}$ и $\vec{h}_k^{(-)}$ – векторные функции полей обычных волн, возбуждаемых в двухслойном волноводе источником; $\vec{\xi}_k^{(+)}$ и $\vec{h}_k^{(+)}$ – векторные функции полей обычных волн, возбуждаемых в двухслойном волноводе на его стыке с однородно заполненным; (+) соответствуют волнам, распространяющимся навстречу оси z , (-) – в направлении оси z .

Дифракционная задача для стыка двух волноводов сводится к выполнению граничных условий при $z = l_1$:

$$\vec{E}_{I\tau} = \vec{E}_{II\tau}; \vec{H}_{I\tau} = \vec{H}_{II\tau}, \quad (4)$$

из которых получается связь между амплитудными коэффициентами (1) и (3). $E_{1,II\tau}$, $H_{1,II\tau}$ – компоненты, тангенциальные к границе.

В результате имеем:

$$R_1 = A_1 e^{-i\beta_1 l_1}; \bar{R}_1 = \bar{A}_1 e^{-i\beta_1 l_1}, \quad (5)$$

а с учетом (2):

$$\begin{aligned} \bar{R}_1 &= -R_1^*; |R_1| = |\bar{R}_1| = |A_1| e^{-\beta_1 l_1}; \\ R_2 &= \chi R_1; \bar{R}_2 = \chi^* \bar{R}_1; \bar{R}_2 = -R_2^*; \\ \chi &= \frac{\alpha_1^2}{\alpha_2^2} J_1(\alpha_1 a). \end{aligned} \quad (6)$$

Таким образом, установили связь между амплитудными коэффициентами комплексных волн, возникающих на стыке двух волноводов. Эта связь имеет место при условии совместного возбуждения источником, расположенным в плоскости $z = 0$, пары комплексных волн, связанных соотношением (2).

Компоненты полей с амплитудными коэффициентами (5) могут быть подставлены в выражение для вектора Умова-Пойтинга, определяющего плотность потока мощности. Энергетические линии – линии, вдоль которых распространяется энергия, т.е. направлен вектор Умова-Пойтинга, могут быть построены на основе численного решения уравнения:

$$\frac{dr}{S_r} = \frac{rd\varphi}{S_\varphi} = \frac{dz}{S_z},$$

где S_r , S_φ , S_z – составляющие вектора Умова-Пойтинга.

Энергетические линии, начинающиеся в произвольных точках поперечного сечения волновода, имеют вид пространственных (не лежащих в одной плоскости) кривых. Их построение наглядно демонстрирует распределенный разворот мощности. Действительно, природа собственных КВ объясняется распределенным разворотом потока мощности, в результате которого обратный поток мощности замыкается через источник и влияет на него. Таким образом, распространение энергии в волноводе имеет вихревой характер.

Проведенное рассмотрение показало, что существование собственных комплексных волн, описываемых наиболее общими решениями краевой несамосопряженной задачи, связано с вихревыми потоками энергии, а не с перекачкой энергии между связанными волнами.

Библиографический список

1. Веселов Г.И., Раевский С.Б. Слоистые металлодиэлектрические волноводы. – М.: Радио и связь, 1988.
2. Раевский С.Б. Комплексные волны в двухслойном круглом экранированном волноводе. – Изв. вузов СССР. Сер. Радиофизика, т.15, №1, 1972, с. 112-116.
3. Веселов Г.И., Любимов Л.А. К теории двухслойного диэлектрического волновода в цилиндрическом экране. – Радиотехника и электроника, 1963, т.8, №9, с. 1530-1541.
4. Clarricoats, P.J.B. and Taylor, B.C. Evanescent and Propagating Modes of Dielectric-Loaded Circular Waveguide, Proc. IEEE, 1964, vol.111, no. T-6, pp. 1951-1956.
5. Веселов Г.И., Раевский С.Б. Комплексные волны в поперечно-неоднородных направляющих структурах. – Радиотехника, 1987, т.42, №8, с. 64-67.
6. Раевский А.С., Раевский С.Б. Комплексные волны. – М.: Радиотехника, 2010, 223 с.
7. Раевский А.С., Раевский С.Б. Неоднородные направляющие структуры, описываемые несамосопряженными операторами. – М.: Радиотехника, 110 с.

S.A. Kapustin, V.A. Malakhov, I.V. Malakhova, A.S. Raevsky

**RESEARCH OF THE PHENOMENON OF COMPLEX RESONANCE
USING POWER LINES**

Nizhny Novgorod state technical university n.a. R.E. Alekseev

Purpose: Clarification of the properties of complex resonance arising from the interaction of two complex-conjugate complex waves.

Design/methodology/approach: It is proposed to use energy lines to clarify the properties of complex resonance. Energy lines are lines along which energy propagates, i.e., direction of the Umov-Poyting vector can be constructed based on the numerical solution of the energy line equation.

Finding (Results): The nature of the complex eigenwaves is explained by the distributed reversal of the power flow, as a result of which the reverse power flow is closed through the source and affects it. Thus, the propagation of energy in the waveguide has a vortex character. The performed consideration showed that the existence of complex eigenwaves described by the most general solutions of the boundary-value non-self-adjoint problem is associated with vortex energy flows, and not with the transfer of energy between the coupled waves.

Keywords: complex waves, complex resonance, energy lines

ИССЛЕДОВАНИЕ ПОЛНОГО СПЕКТРА ВОЛН ОТКРЫТОГО ДИЭЛЕКТРИЧЕСКОГО ВОЛНОВОДА С ИСПОЛЬЗОВАНИЕМ ЭНЕРГЕТИЧЕСКИХ ЛИНИЙ

Нижегородский государственный технический университет им. Р.Е. Алексеева

Круглый открытый диэлектрический волновод (ОДВ) является одной из наиболее широко используемых открытых направляющих структур как СВЧ и КВЧ, так и оптического диапазонов волн. Интерес к нему определяется, с одной стороны, широкими возможностями его использования при построении линий связи и функциональных узлов указанных диапазонов, с другой стороны, тем, что ОДВ удобен для теоретического исследования особенностей открытых направляющих структур в силу возможности строгого решения соответствующей ему краевой задачи. Для круглого ОДВ возможна точная запись (в замкнутой форме) дисперсионных уравнений, гарантирующая адекватность используемой математической модели реальной направляющей структуре. Математическая модель круглого регулярного ОДВ позволяет выявить основные особенности волн, направляемых открытыми диэлектрическими волноводами.

Рассмотрен метод исследования полного спектра волн ОДВ с использованием энергетических линий поля. Характер энергетических линий поля волны ОДВ позволяет определить ее принадлежность к конкретному типу волн и объяснить ее физические свойства.

Ключевые слова: дисперсионное уравнение, комплексные волны, открытый диэлектрический волновод, несамосопряженная краевая задача, расчет энергетических линий поля.

Введение

В большинстве работ, посвященных исследованию ОДВ, он рассматривался в основном как структура направляющая поверхностные волны, поля которых экспоненциально убывают в радиальном направлении. Теория именно таких волн получила наиболее полное развитие. Поверхностные волны – основной тип волн, на которых осуществляется перенос энергии в волоконных световодах, для исследования которых используется математическая модель ОДВ. Однако, в целом ряде устройств (это, в первую очередь, относится к плавным переходам линий передач с малым замедлением и к антеннам не осевого излучения) поверхностные волны используются лишь для возбуждения, подвергаясь затем преобразованию в волны, слабо связанные с направляющей структурой, имеющие большую поперечную протяженность поля. При этом возникают задачи о расчете поля излучения и характеристик передачи нерегулярных участков тракта, при решении которых неизбежно приходится учитывать полный спектр волн диэлектрического волновода.

В отличие от экранированных направляющих структур полный спектр собственных значений краевых задач для открытых волнопроводов, помимо дискретной, имеет (в силу отсутствия на внешней границе краевого условия, соответствующего задаче Штурма-Лиувилля), непрерывную часть, представляемую интегралом по одному из волновых чисел. Волновые числа непрерывного спектра изменяются в пределах от нуля до бесконечности и соответствуют полю излучения из волновода и полю вблизи источника. Эти поля частично описываются также различными типами комплексных волн (КВ), входящих в дискретную часть спектра. Комплексность волновых чисел указанных волн не связана с тепловыми потерями в средах направляющей структуры.

Одним из способов изучения свойств ОДВ является исследование ее полного спектра волн с использованием энергетических линий поля.

Дисперсионная задача для волн открытого диэлектрического волновода

Вид комплексных собственных значений зависит от особенностей характеристических (дисперсионных) уравнений. Дисперсионные уравнения (ДУ) диссипативных электродинамических структур всегда комплексные (их решения являются комплексными величинами) потому, что диссипативные структуры описываются несамосопряженными операторами, поскольку для них не выполняется первое условие самосопряженности [1-3]. Для структур без диссипации энергии самосопряженность определяется [1-3] граничными условиями. Рассмотрим особенности решений ДУ волн, описываемых несамосопряженными операторами на примере ОДВ.

ДУ волн ОДВ имеет вид:

$$\left[\frac{\varepsilon_1 J'_n(\alpha_1 R)}{\alpha_1^2 J_n(\alpha_1 R)} - \frac{\varepsilon_2 H_n^{(2)'}(\alpha_2 R)}{\alpha_2 H_n^{(2)}(\alpha_2 R)} \right] * \left[\frac{\mu_1 J'_n(\alpha_1 R)}{\alpha_1 J_n(\alpha_1 R)} - \frac{\mu_2 H_n^{(2)'}(\alpha_2 R)}{\alpha_2 H_n^{(2)}(\alpha_2 R)} \right] = \frac{n^2(\alpha_1^2 - \alpha_2^2)}{R^2 \alpha_1^4 \alpha_2^4} * (\varepsilon_2 \mu_2 \alpha_1^2 - \varepsilon_1 \mu_1 \alpha_2^2), \quad (1)$$

где R – радиус внутреннего слоя ОДВ; $\alpha_{1,2}$ – поперечные волновые числа для внутреннего и наружного слоев; ε_1, μ_1 и ε_2, μ_2 – параметры слоев;

$J_n(\alpha_2 r), H_n^{(2)}(\alpha_2 r)$ – цилиндрические функции 1-го и 2-го рода.

Поперечные волновые числа $\alpha_{1,2} = \gamma_{1,2} + i\delta_{1,2}$ связаны с продольным волновым числом $\beta = \beta_1 + i\beta_2$ соотношениями:

$$\alpha_{1,2}^2 = \beta^2 + \varepsilon_{1,2} \mu_{1,2} \omega^2. \quad (2)$$

Таким образом, дисперсионная задача состоит из трех уравнений: (1), (2) относительно трех неизвестных: $\alpha_1, \alpha_2, \beta$. При действительных $\varepsilon_{1,2}$ и $\mu_{1,2}$ комплексность ДУ полностью определяется комплексностью волновых чисел (2). Отсюда можно сделать вывод, что дисперсионная задача имеет пары комплексно-сопряженных решений.

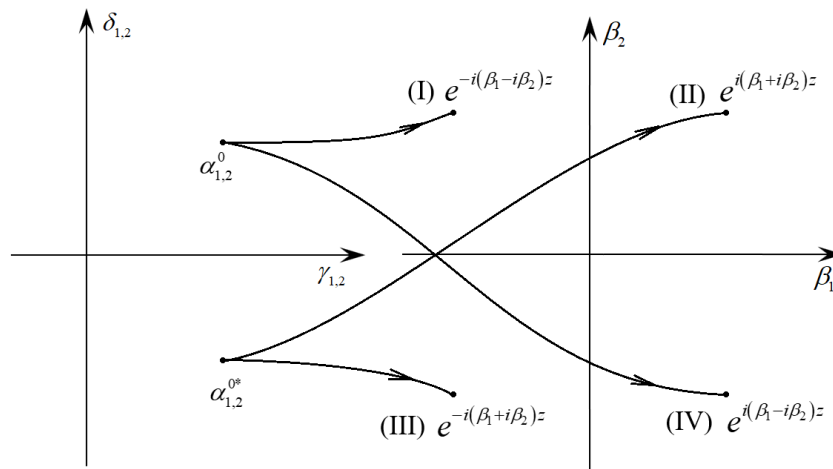


Рис. 1. Решения дисперсионных уравнений, соответствующие парам сопряженных комплексных волн

Действительно, уравнения (1) и (2) полностью определены на комплексных плоскостях: $\alpha_1, \alpha_2, \beta$.

Комплексно-сопряженные решения ДУ дают две пары волн:

$$\begin{aligned} & e^{-i(\beta_1 - i\beta_2)z}, e^{i(\beta_1 + i\beta_2)z}; \quad (\text{I}; \text{II}) \\ & e^{-i(\beta_1 + i\beta_2)z}, e^{i(\beta_1 - i\beta_2)z}. \quad (\text{III}; \text{IV}) \end{aligned} \quad (3)$$

Волны I, II распространяются справа от источника, расположенного при $z = 0$, волны III, IV – слева.

В том случае, когда параметры: $\varepsilon_{1,2}$ и $\mu_{1,2}$ являются комплексными величинами, то есть структуры диссипативные, ДУ не имеют комплексно-сопряженных решений. При этом волны: I и II, III и IV существуют независимо. В недиссипативных структурах любые источники создают волны: I и II, III и IV с комплексно-сопряженными волновыми

числами и амплитудами. Поскольку существование таких пар волн обеспечивается источником, их можно классифицировать как присоединенные к последнему. Вместе эти пары комплексных волн образуют колебание, присоединенное к источнику, которое создает комплексный резонанс. Совместно существующие волны: I и II, III и IV создают, как следует из (3), поле, локализованное вблизи источника, соответствующее комплексному резонансу.

Исследование полного спектра волн ОДВ с использованием энергетических линий поля

В общем случае, в ОДВ могут существовать следующие типы комплексных волн:

- гибридные комплексные волны, как с экспоненциально убывающим, так и с экспоненциально нарастающим (вытекающие волны) по радиальной координате полем
- несобственные комплексные волны.

Полный спектр волн ОДВ может быть исследован путем изучения трансформации их структуры полей во всем частотном диапазоне. Для этого решается дисперсионная задача (1), (2), находятся комплексные числа $\alpha_1, \alpha_2, \beta$.

После этого один из амплитудных коэффициентов $A_{1,2}, B_{1,2}$ векторов Герца приравняется 1 и относительно него рассчитываются оставшиеся амплитудные коэффициенты для конкретных значений $\alpha_1, \alpha_2, \beta$.

Вектора Герца для двух областей ОДВ записываются следующим образом:

I область:

$$\begin{aligned} \Pi_z^e &= A_1 I_n(\alpha_1 r) \cos(n\varphi) e^{-i\beta z}, \\ \Pi_z^m &= B_1 I_n(\alpha_1 r) \sin(n\varphi) e^{-i\beta z}. \end{aligned} \quad (4)$$

II область:

$$\begin{aligned} \Pi_z^e &= A_2 H_n^{(2)}(\alpha_2 r) \cos(n\varphi) e^{-i\beta z}, \\ \Pi_z^m &= B_2 H_n^{(2)}(\alpha_2 r) \sin(n\varphi) e^{-i\beta z}. \end{aligned} \quad (5)$$

Поперечный поток мощности, определяемый комплексным вектором Умова-Пойнтинга, записывается следующим образом:

$$\vec{S}_\perp^{ik} = S_r^{ik} * \vec{e}_1 + S_\varphi^{ik} * \vec{e}_2,$$

где $\vec{e}_{1,2}$ – соответствующие единичные орты

$$\begin{aligned} S_\varphi^{ik} &= \frac{1}{2} \operatorname{Re}(E_\varphi H_z^* - E_z H_\varphi^*), \\ S_r^{ik} &= \frac{1}{2} \operatorname{Re}(E_z H_r^* - E_r H_z^*). \end{aligned} \quad (6)$$

Продольный поток мощности определяется следующим образом:

$$S_z^{ik} = \frac{1}{2} \operatorname{Re}(E_r H_\varphi^* - E_\varphi H_r^*). \quad (7)$$

Компоненты полей для 2-х областей записываются следующим образом:

I область:

$$\begin{aligned} E_z &= \alpha_1^2 A_1 I_n(\alpha_1 r) \cos(n\varphi) e^{-i\beta z}, \\ H_z &= \alpha_1^2 B_1 I_n(\alpha_1 r) \sin(n\varphi) e^{-i\beta z}, \\ E_r &= -i \left[\beta \alpha_1 A_1 I_n'(\alpha_1 r) \cos(n\varphi) + \frac{\omega \mu_1}{r} n B_1 I_n(\alpha_1 r) \cos(n\varphi) \right] e^{-i\beta z}, \\ H_r &= -i \left[\frac{\omega \varepsilon_1}{r} n A_1 I_n(\alpha_1 r) \sin(n\varphi) + \beta \alpha_1 A_1^m I_n'(\alpha_1 r) \sin(n\varphi) \right] e^{-i\beta z}, \\ E_\varphi &= i \left[\frac{\beta n}{r} A_1(\alpha_1 r) \sin(n\varphi) + \omega \mu_1 \alpha_1 B_1 I_n(\alpha_1 r) \sin(n\varphi) \right] e^{-i\beta z}, \\ H_\varphi &= -i \left[\omega \varepsilon_1 \alpha_1 A_1 I_n'(\alpha_1 r) \cos(n\varphi) + \frac{\beta n}{r} A_1^m(\alpha_1 r) \cos(n\varphi) \right] e^{i\beta z}. \end{aligned} \quad (8)$$

II область:

$$\begin{aligned}
 E_z &= \alpha_2^2 A_2 H_n^{(2)}(\alpha_2 r) \cos(n\varphi) e^{-i\beta z}, \\
 H_z &= \alpha_2^2 B_2 H_n^{(2)}(\alpha_2 r) \sin(n\varphi) e^{-i\beta z}, \\
 E_r &= -i \left[\beta \alpha_2 A_2 H_n^{(2)'}(\alpha_2 r) \cos(n\varphi) + \frac{\omega \mu_2}{r} n B_2 I_n(\alpha_2 r) \cos(n\varphi) \right] e^{-i\beta z}, \\
 H_r &= -i \left[\frac{\omega \varepsilon_2 n}{r} A_2 H_n^{(2)}(\alpha_2 r) + \beta \alpha_2 B_2 H_n^{(2)'}(\alpha_2 r) \right] \sin(n\varphi) e^{-i\beta z}, \\
 E_\varphi &= i \left[\frac{\beta n}{r} A_2 H_n^{(2)}(\alpha_2 r) + \omega \mu_2 \alpha_2 B_2 H_n^{(2)'}(\alpha_2 r) \right] \sin(n\varphi) e^{-i\beta z}, \\
 H_\varphi &= i \left[\omega \varepsilon_2 \alpha_2 A_2 H_n^{(2)'}(\alpha_2 r) + \frac{\beta n}{r} B_2 H_n^{(2)}(\alpha_2 r) \right] \cos(n\varphi) e^{-i\beta z}.
 \end{aligned} \tag{9}$$

Расчет одной энергетической линии поля волны производится следующим образом:

1. Для фиксированных значений $\alpha_1, \alpha_2, \beta$ находятся амплитудные коэффициенты $A_{1,2}, B_{1,2}$ при соответствующих им векторах Герца.

2. Записываются начальные координаты z_0, φ_0, r_0 для соответствующей области КДВ, внутренней или внешней. Например, для внутренней области стандартные значения могут быть: $z_0 = 0, \varphi_0 = 0, r_0 \in [0, R]$.

3. Производится расчет координаты следующей точки с использованием формул (6)-(9):

$$z = z_0 + S_z^{ik} * e_z, \quad \varphi = \varphi_0 + S_\varphi^{ik} * e_\varphi, \quad r = r_0 + S_r^{ik} * e_r. \tag{10}$$

При этом значения $e_{z,\varphi,r}$ подбираются таким образом, чтобы энергетические линии поля менялись плавно [3].

4. Расчет производится последовательно до тех пор, пока не выполнится одно из условий: $(r = r_0) \parallel (r = R), z = z_0, (\varphi = \varphi_0) \parallel (\varphi = \frac{\pi}{2})$.

5. Если выполнилось условие $(r = R)$, производится переход из одной области КДВ в другую, путем замены соответствующих компонент поля в уравнениях (6) и (7). После этого переходим к шагу 3 с сохранением текущих координат.

Количество энергетических линий поля, подлежащих расчету, определяется опытным путем. После окончания расчетов все линии поля наносятся на график и анализируются совместно. Затем расчет повторяется для другой частотной области дисперсионного уравнения.

Следовательно, используя метод расчета энергетических линий поля волны можно провести анализ полного спектра волн ОДВ.

Заключение

Для любой комплексной волны ОДВ, используя формулы (10) и зная значения амплитудных коэффициентов $A_{1,2}, B_{1,2}$ и волновых чисел $\alpha_1, \alpha_2, \beta$, можно построить энергетические линии поля, вдоль которых переносится энергия данной волны. Характер энергетических линий поля КВ позволяет определить ее принадлежность к конкретному типу волн и объяснить ее физические свойства. Таким образом, используя описанный метод расчета энергетических линий поля можно исследовать полный спектр волн ОДВ.

Библиографический список

1. Раевский А.С. Комплексные волны / А.С. Раевский, С.Б. Раевский. – М.: Радиотехника, 2010, с. 223.
2. Раевский А.С. Неоднородные направляющие структуры, описываемые несамосопряженными операторами / А.С. Раевский, С.Б. Раевский. – М.: Радиотехника, с. 110.
3. Веселов Г.И. Слоистые металло-диэлектрические волноводы / Г.И. Веселов, С.Б. Раевский. – М.: Радио и связь, 1988.

A.V. Lvutin, A.S. Raevskii

INVESTIGATION OF THE FULL SPECTRUM OF WAVES OF AN OPEN DIELECTRIC WAVEGUIDE USING ENERGY LINES

Purpose: investigation of the full spectrum of waves, including leaky waves, of an open dielectric waveguide.

Design/methodology/approach: solution of the dispersion equation of an open dielectric waveguide and further calculation and analysis of the field energy lines

Findings: The use of the method of calculation the energy lines of the field allows one to analyze the physical properties of waves in an open dielectric waveguide.

Originality/value: The presented work describes in detail the method for the calculating the energy lines of the field of the entire wave spectrum of an open dielectric waveguide.

Keywords: dispersion equation, complex waves, open dielectric waveguide, non-self-adjoint boundary value problem, calculation of the field energy lines/

ОБ ОСОБЕННОСТЯХ РЕШЕНИЯ ЗАДАЧИ О ВОЗБУЖДЕНИИ ОТКРЫТОГО ДИЭЛЕКТРИЧЕСКОГО ВОЛНОВОДА С ИСПОЛЬЗОВАНИЕМ ИЗВЕСТНЫХ САПР

Нижегородский государственный технический университет им. Р.Е. Алексеева

В представленной работе предложены варианты конструкций схем возбуждения открытого диэлектрического волновода. Исследование проводится с применением современных инструментов электродинамического моделирования и анализа в системе автоматического проектирования (САПР).

Эффективное возбуждение вытекающих и поверхностных волн позволяет решать актуальные задачи в области волноводных и излучающих структур.

Ключевые слова: антенна, вытекающие волны, поверхностные волны, открытый диэлектрический волновод, САПР.

Введение

Обеспечение технологического преимущества отечественных систем радиолокации и связи невозможно без разработки и внедрения новых перспективных устройств. Накопленные за долгие годы исследований эмпирические данные об антенных устройствах на основе открытых диэлектрических волноводов (ОДВ) лежат в основе многих эффективных методов их возбуждения. В представленной работе рассмотрены вопросы, связанные со спецификой возбуждения вытекающих и поверхностных волн ОДВ, а также эффективный метод электродинамического анализа характеристик таких систем в зарубежных САПР.

Цель исследования – анализ методов возбуждения вытекающих и поверхностных волн открытых диэлектрических волноводов на примере круглого ОДВ с целью использования результатов в работах по созданию отечественного САПР ЛОГОС ЭМИ.

Излучающие системы на основе ОДВ

Одним из широко применяемых типов электродинамических направляющих высокочастотных структур является круглый ОДВ. Для построения излучающих систем зачастую используют режим работы на основной волне HE_{11} , которая является поверхностной волной, её поле быстрее, чем экспоненциально, убывает в радиальном направлении при удалении от волновода [1].

Чтобы обеспечить одномодовый режим на волне HE_{11} , диаметр ОДВ должен удовлетворять условию:

$$d < \frac{\lambda_0}{1.3\sqrt{\varepsilon_1 - \varepsilon_2}} \quad (1)$$

Фазовая скорость, с которой электромагнитная волна распространяется вдоль ОДВ, зависит от диэлектрической проницаемости материала, диаметра стержня и частоты. Величина диаметра стержня, при котором фазовая скорость практически сравнивается со скоростью света в диэлектрике, может быть определена по формуле [2]:

$$d = \frac{0.4\lambda_0}{\sqrt{\varepsilon_1 - \varepsilon_2}} \quad (2)$$

Выбор длины ОДВ выбирается из условия максимального КНД антенны и из теории антенн бегущей волны, может быть определен как:

$$L_{omn} = \frac{\lambda}{2\left(\frac{1}{\gamma} - 1\right)}, \quad (3)$$

где $\gamma \sim 0,9$ (отношение фазовой скорости к скорости света).

Если в качестве излучателя системы выбрать ОДВ с не изменяющей по длине цилиндрической формой, то значительная часть энергии будет отражаться от его конца в обратном направлении, что приведет к увеличению уровня боковых и задних лепестков диаграммы направленности. Существует несколько методов улучшения согласования диэлектрического волновода со свободным пространством, позволяющих улучшить характеристики. Для согласования открытого конца ОДВ его сужают либо в виде конической формы, либо в виде скругления на конце. Часто для подведения энергии к диэлектрической антенне используется полый металлический волновод. Согласование стержня с волноводом может быть осуществлено расположением на конце волновода согласующей камеры, внутри которой начальный участок ОДВ также срезается на конус.

Система возбуждения ОДВ от штыревого вибратора

Как один из конструктивных вариантов антенны на основе ОДВ может быть рассмотрена система, состоящая из фидерной части, устройства возбуждения, участка сопряжения и устройства излучения (рис. 1).

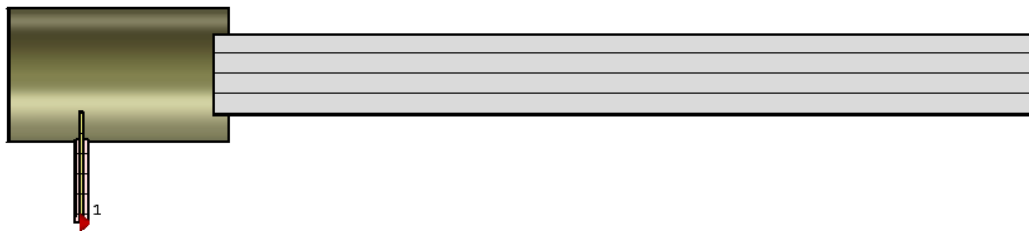


Рис. 1. Конструкция возбуждения ОДВ от штыревого вибратора

Электромагнитная энергия подводится к ОДВ от фидера в виде коаксиальной линии, внутренний проводник которой продолжается, образуя возбуждающую систему – штыревой вибратор, возбуждающий основной тип волны H_{11} в экранированном волноводе круглого сечения.

Питающий коаксиальный кабель характеризуется радиусами внешнего и внутреннего проводников, а также диэлектрической проницаемостью заполняющего материала и выбирается исходя из потребности обеспечения волнового сопротивления фидера на входе антенны.

Для согласования волнового сопротивления коаксиального кабеля с входным сопротивлением антенны необходимо найти нужную величину действующей высоты возбудителя (штыря), при которой эти сопротивления равны. Расстояние от закорачивающей стенки до оси штыря выбирается равным $\lambda_v/4$, где λ_v – длина волны в волноводе с волной H_{11} .

$$\lambda_g \cong \frac{\lambda W_{h11}}{120\pi}, \quad (4)$$

где W_{h11} – волновое сопротивление круглого волновода, для волны H_{11} .

Определить наиболее оптимальный радиус экранированного волновода по соотношениям для определения критических длин волн H_{11} и E_{01} :

$$\lambda_{кр}(H_{11}) \cong 3,41a \text{ и } \lambda_{кр}(E_{01}) \cong 2,61a. \quad (5)$$

В режиме осевого излучения максимум диаграммы направленности расположен в направлении оси диэлектрического стержня, когда основное влияние на диаграмму направленности оказывает поверхностная волна HE_{11} круглого ОДВ [3]. Анализ антенн с использованием ОДВ возможен в современных системах автоматического проектирования, позволяющих с высокой степенью точности проводить расчеты и оптимизировать систему для достижения требуемых характеристик. На рис. 2 представлена диаграмма направленности модели, построенной на основе описанных выше рекомендаций для подбора размеров для частоты $F=2,45$ ГГц. В качестве материала ОДВ выбран полистирол, относительная диэлектрическая проницаемость которого равна 2,5.

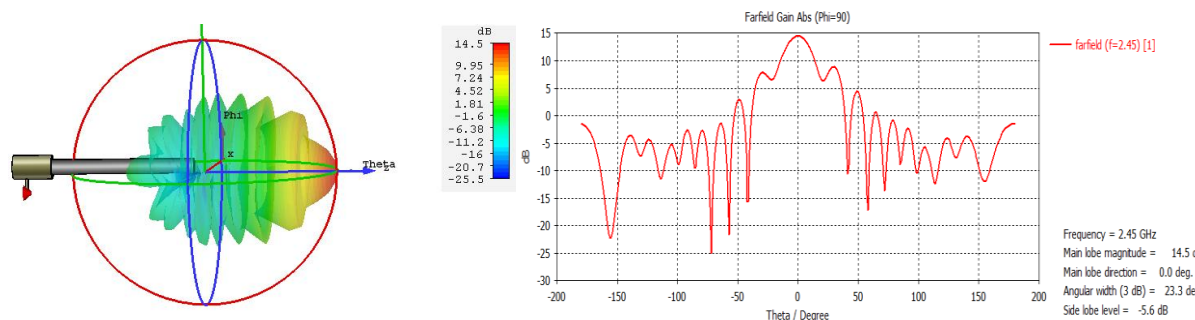


Рис. 2. Диаграмма направленности при возбуждении ОДВ от штыревого вибратора

Величина коэффициента усиления антенны, уровень боковых лепестков и обратного излучения могут быть улучшены за счет изменения формы ОДВ, а также подбором дополнительных согласующих элементов.

Система осевого возбуждения ОДВ от коаксиальной линии

Другой вариант возбуждения – осевое от коаксиальной линии (рис. 3).

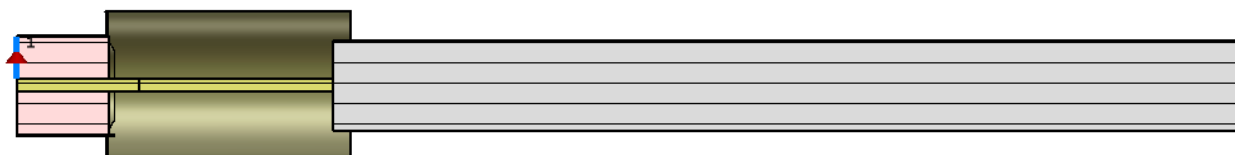


Рис. 3. Конструкция осевого возбуждения ОДВ от коаксиальной линии

Для определения размеров коаксиальной линии, в которой распространяется только основной тип волны T ($\lambda_{крT} = \infty$), можно воспользоваться формулой для определения критической длины первой из волн высшего типа квази- H_{11} :

$$\lambda_{кр}^{H_{11}} \approx \pi(R_1 + R_2), \quad (6)$$

где R_1 и R_2 – радиусы внутреннего и внешнего проводников коаксиального кабеля.

На рис. 4 представлена диаграмма направленности диэлектрической антенны при осевом возбуждении ОДВ от коаксиальной линии. Анализ проводился для модели с теми же параметрами диэлектрического стержня и на той же частоте, что и в предыдущем рассмотрении для случая штыревого возбуждения.

Сравнивая такой способ возбуждения ОДВ с предыдущим, можем видеть возрастание уровня боковых лепестков и обратного излучения, а также смещение от оси ОДВ максимума диаграммы направленности. Это можно объяснить тем, что при данном способе возбуждения наряду с основной волной HE_{11} в отрезке ОДВ возбуждается волна E_{01} , поле которой по структуре сходно с полем T -волны коаксиального кабеля.

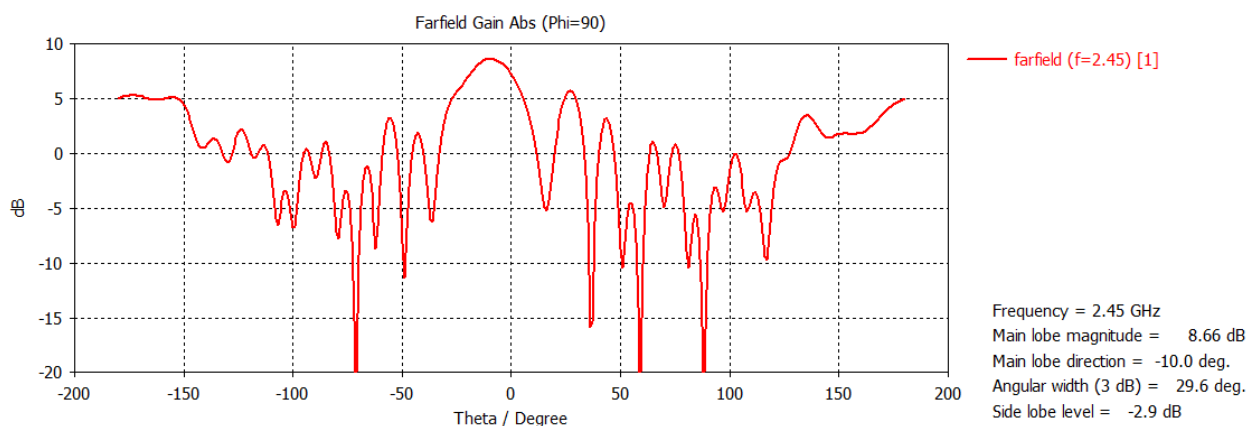


Рис. 4. Диаграмма направленности при возбуждении ОДВ от коаксиальной линии

Выводы

Использование современных САПР позволяет успешно рассчитывать характеристики антенно-фидерных систем, в частности, антенн на основе ОДВ. Вопрос согласования ОДВ со свободным пространством и фидерной системой может быть решен посредством построения моделей и их последовательной оптимизации для достижения требуемых характеристик антенн. В настоящее время в связи с тенденцией импортозамещения решается задача создания отечественного САПР ЛОГОС ЭМИ – программы трёхмерного электромагнитного моделирования. Создателям данной программы для обеспечения её конкурентоспособности необходимы результаты анализа возможностей зарубежных САПР, в частности, в плане задания способов возбуждения открытых диэлектрических волноводов.

Библиографический список

1. Унгер Х.Г. Планарные и волоконные оптические волноводы // Х.Г. Унгер. – М.: Мир, 1988. – 656 с.
2. Дмитриев Д.Д. Устройства СВЧ и антенны. – 2019.
3. Васильев Е. Н. Возбуждение тел вращения //М.: Радио и связь. – 1987. – Т. 270.

M.S. Nechaeva, A.S. Raevskii

SPECIFIC FEATURES OF THE SOLUTION TO THE PROBLEM OF EXCITATION OF AN OPEN DIELECTRIC WAVEGUIDE USING KNOWN CAD

Nizhny Novgorod state technical university n.a.R.E. Alekseeva

Purpose: Investigation of electrodynamic structures for the excitation of leaky and surface waves of a circular open dielectric waveguide

Design/methodology/approach: CAD analysis of emitting systems of leaky and surface waves of an open dielectric waveguide excited from coaxial and circular shielded waveguides

Findings: The use of modern CAD makes it possible to successfully analyze the characteristics of antennas based on open dielectric waveguides and their excitation systems for the effective solution of urgent problems.

Originality/value: In the presented work, schemes of effective excitation of various types of waves of an open dielectric waveguide and the method of electrodynamic analysis of the characteristics and tuning of such systems in CAD are considered.

Keywords: antenna, leaky waves, surface waves, open dielectric waveguide, CAD.

ИССЛЕДОВАНИЕ СПЕКТРАЛЬНЫХ ХАРАКТЕРИСТИК ДВУХКОНТУРНОГО ОПТОЭЛЕКТРОННОГО ГЕНЕРАТОРА

Нижегородский государственный технический университет им. Р.Е. Алексеева

В данной работе рассматривается реализация двухконтурного оптоэлектронного генератора (ОЭГ) радиочастотных сигналов; исследуется уровень спектральной плотности мощности фазовых флуктуаций, а также спектральные характеристики одноконтурного и двухконтурного генератора.

Ключевые слова: оптоэлектронный генератор, радиофотоника, оптическая линия задержки, фазовый шум, автогенератор.

Введение

В области коммуникаций для передачи больших объемов данных используется цифровая обработка сигналов, однако существует ряд технических задач, для решения которых требуются не цифровые, а именно аналоговые системы. Возможность обработки сигнала в аналоговой среде может упростить общее проектирование, особенно в широкополосных системах, где затруднительно обеспечить требуемую полосу пропускания, используя цифровые средства.

В частности, актуальным вопросом является разработка радиочастотных генераторов, работающих в сверхвысокочастотном диапазоне, обладающей высокой стабильностью частоты, широкой полосой перестройки, а также небольшими размерами и массой. Такие генераторы сигналов СВЧ диапазона применяются в составе радиоэлектронных приборов и устройств телекоммуникационного, радионавигационного и метрологического назначений, поэтому совершенствование их основных характеристик является важной целью для дальнейшего развития радиоэлектронных устройств.

Технологичной особенностью оптоэлектронных генераторов (ОЭГ) является совокупное использование твердотельных компонентов современной ВЧ и СВЧ оптоэлектроники, волоконных световодов и традиционной элементной базы СВЧ/КВЧ техники. Важным преимуществом оптоэлектронных генераторов, в сравнении с генераторами, построенными по классической схемотехнике, выступает отсутствие компромисса между полосой перестройки частоты генерации прибора и уровнем частотных шумов.

Исследование спектральных характеристик

По принципу работы и схеме построения ОЭГ аналогичен классическому радиочастотному автогенератору с линией задержки в цепи обратной связи. Особенностью его работы является многочастотный характер генерируемых колебаний, для которых выполняются условия баланса амплитуд и баланса фаз [1].

В случае ОЭГ кольцо обратной связи с резонансной системой и элементом задержки выполняется с помощью одномодового волоконного световода определенной длины. Время задержки сигнала в петле обратной связи является значимым параметром оптоэлектронного генератора, влияющим на шумовые и спектральные характеристики генератора.

При увеличении длины волоконно-оптического тракта увеличивается время задержки в цепи обратной связи, уменьшая тем самым уровень частотных шумов. Однако, при увеличении длины ВОТ также уменьшается и расстояние между соседними модами генерации ОЭГ, ужесточая требования к полосе пропускания ППФ. Решением этой проблемы стало использование многоконтурной схемы ОЭГ.

Для практической реализации в рамках имеющейся элементной базы была выбрана схема двухконтурного ОЭГ, представленная на рис. 1.

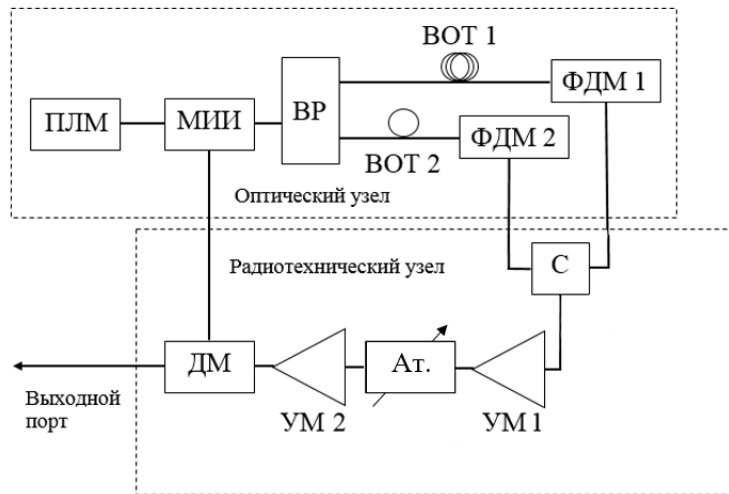


Рис. 1. Структурная схема реализуемого ОЭГ

В макете в качестве волоконно-оптических трактов используются волоконные световоды длинами $l_1 = 1000$ м и $l_2 = 30$ м. Произведем расчет характеристик предлагаемого двухконтурного оптоэлектронного генератора.

Определим время задержки распространения сигнала в ВОТ. Показатель преломления кварцевого световода $n = 1,45$, скорость света в вакууме $c = 3 \cdot 10^8$ м/с. Тогда

$$\tau_1 = \frac{n \cdot l_1}{c} = \frac{1,45 \cdot 1000}{3 \cdot 10^8} = 4,83 \cdot 10^{-6} \text{ с}, \tau_2 = \frac{n \cdot l_2}{c} = \frac{1,45 \cdot 30}{3 \cdot 10^8} = 1,45 \cdot 10^{-7} \text{ с}.$$

Расстояние между соседними модами при включении только ВОТ1 и только ВОТ2 определяется следующим образом:

$$\Delta f_1 = \frac{1}{\tau_1} = \frac{1}{4,83 \cdot 10^{-6}} \approx 207 \text{ кГц}, \quad \Delta f_2 = \frac{1}{\tau_2} = \frac{1}{1,45 \cdot 10^{-7}} \approx 6,9 \text{ МГц},$$

Так как в двухконтурном оптоэлектронном генераторе область свободной дисперсии определяется контуром меньшей длины, то $\Delta f = \Delta f_2 \approx 6,9 \text{ МГц}$.

После сборки макета исследуемого генератора с помощью анализатора спектра Rohde&Schwarz FSU50 был исследован спектр выходного сигнала генератора. На рисунках 2,а и 2,б изображены спектры генерации одноконтурного оптоэлектронного генератора с длинами волоконных трактов $l_1 = 1000$ м и $l_2 = 30$ м соответственно.

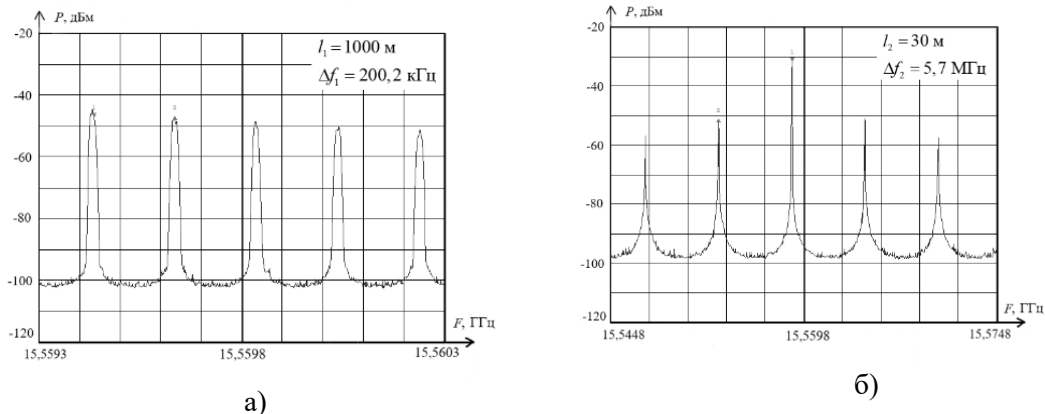


Рис. 2. Спектр одноконтурного ОЭГ при длине волоконной линии 1000 м (а) и 30 м (б)

Спектр генерации двухконтурного оптоэлектронного генератора с одновременно подключенными волоконными трактами l_1 и l_2 представлен на рисунке 3. На рисунке видно, что в спектре двухконтурного генератора присутствуют как составляющие, обусловленные короткой волоконной линией, так и составляющие, обусловленные линией большей длины.

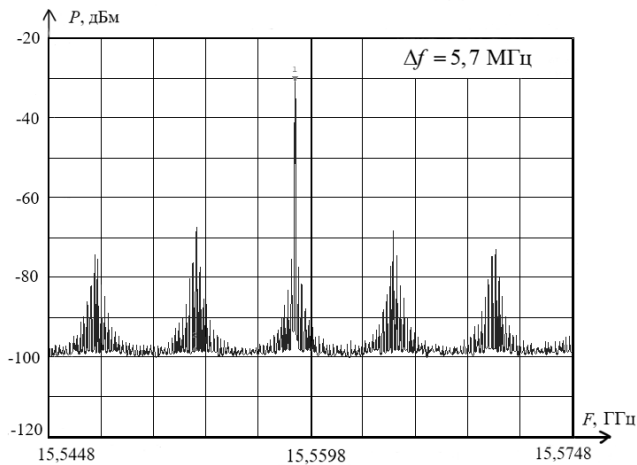


Рис. 3. Спектр генерации ОЭГ при одновременном включении ВОТ1 и ВОТ2

Исследование шумовых характеристик

Фазовый шум является важным параметром в оценке генераторов. Генераторы, применяемые в качестве опорных и измерительных, непременно должны обладать высокой стабильностью частоты генерируемого сигнала. Колебания в автогенераторах возникают благодаря воздействию собственных шумов в элементах схемы, что является одной из причин возникновения шумов в спектре выходного сигнала генератора. К собственным шумам можно отнести тепловые, дробовые и фликкер-шумы в полупроводниковых элементах генераторов [2].

Для количественной оценки фазовых шумов обычно используется его спектральная плотность мощности (СПМ) в одиночной боковой полосе $L(f)$ при заданной отстройке f' от несущей частоты f_0 генератора. Снижение этой величины в генераторе достигается за счет увеличения добротности Q входящего в его состав резонатора. В оптоэлектронном генераторе используются оптоволоконные линии задержки, их эквивалентная нагруженная добротность определяется формулой

$$Q = Q_{\text{опт}} \cdot \frac{f_0}{f_{\text{опт}}}, \quad (1)$$

где $Q_{\text{опт}}$ – эквивалентная добротность волоконной линии на оптической частоте $f_{\text{опт}}$,

$$Q_{\text{опт}} = 2\pi f_{\text{опт}} \tau. \quad (2)$$

Для оценки спектральной плотности мощности фазового шума будем использовать шумовую модель, основанную на формуле Лисона для генератора с обратной связью [3]

$$L(f') = 10 \lg \left\{ \delta \left[\left(1 + \frac{1}{f'^2} \left(\frac{f_0}{2Q} \right)^2 \right) \left(1 + \frac{f_\alpha}{f'} \right) \right] \right\}, \quad (3)$$

где f_α – верхняя граничная частота фликкер-шумов активного прибора;

δ – отношение шум/сигнал оптоэлектронного генератора, определяемое выражением

$$\delta = \frac{P_N G_A^2}{P} = \frac{[4k_b T N_f + 2e I_\Phi R + N_{RIN} I_{\Phi Д}^2 R]}{P}; \quad (4)$$

P_N – плотность шумов на управляющем входе модулятора;

$G_A = G_{\text{ПЗУ}} G_{\text{УМ}}$ – коэффициент усиления радиотехнического узла ОЭГ;

P – СВЧ мощность, приведенная ко входу ПЭУ;

k_B – постоянная Больцмана;

T – температура окружающей среды;

N_f – коэффициент шума радиотехнического узла ОЭГ;

e – заряд электрона;

I_Φ – величина фототока на сопротивлении нагрузки ФДМ (R);

N_{RIN} – относительный шум интенсивности излучения лазера.

Выражение (3) описывает четыре основные области зависимости спектральной плотности мощности шума от отстройки f' . Однако, в СВЧ-генераторах область с зависимостью вида $1/f$ обычно отсутствует вследствие доминирования шума вида $1/f'^2$.

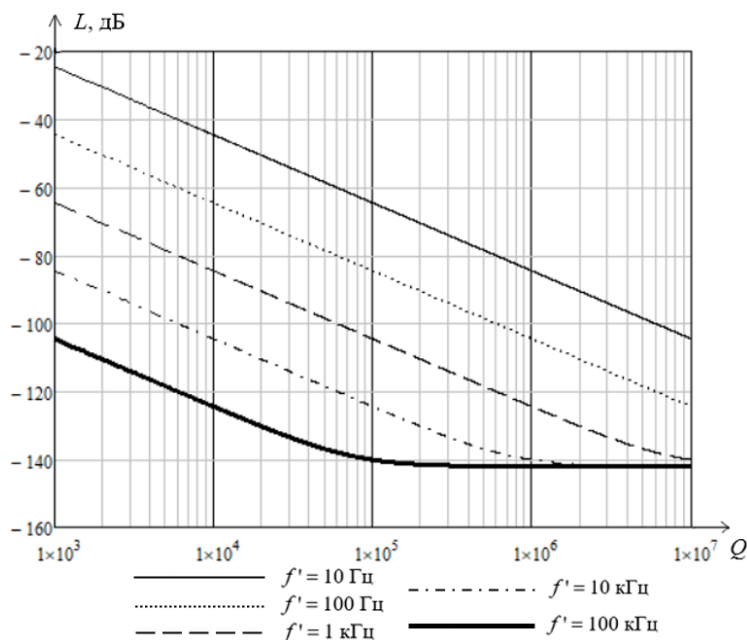


Рис. 4. Зависимость СПМ фазовых шумов от эквивалентной добротности волоконной линии при разных значениях отстройки частоты

Точка, в которой постоянный уровень шума переходит в уровень, нарастающий как 20 дБ на декаду, соответствует полуширине полосы пропускания резонансной системы. Чем выше добротность резонатора, тем ниже частота отстройки, соответствующая переходу на постоянный уровень шумов.

Можно получить зависимость уровня фазовых шумов от эквивалентной добротности волоконной линии при фиксированной частоте отстройки. На рисунке 4 приведена такая зависимость при частотах отстройки $f' = 10$ Гц; 100 Гц; 1 кГц; 10 кГц и 100 кГц. Очевидно, что при большей добротности уровень шумов оказывается ниже. Эквивалентная добротность волоконного тракта пропорциональна его длине, а значит, для достижения лучших шумовых характеристик необходимо увеличивать длину волоконно-оптического тракта.

Рассчитаем спектральную плотность шумов одноконтурного ОЭГ. В соответствии с формулами (1) и (2) эквивалентная нагруженная добротность линии задержки

$$Q = 2\pi f_{\text{опт}} \tau \frac{f_0}{f_{\text{опт}}} = 2\pi f_0 \frac{nl}{c} \quad (5)$$

Для волоконных линий длинами $l_1 = 1000$ м и $l_2 = 30$ м эквивалентная нагруженная добротность составляет

$$Q_1 = 2\pi f_0 \frac{nl_1}{c} = 4,6 \cdot 10^5, \quad Q_2 = 2\pi f_0 \frac{nl_2}{c} = 1,4 \cdot 10^4.$$

У разных устройств точка пересечения белого шума и шума $1/f$ разная, и в имеющемся активном элементе величина f_α граничной фликкерной частоты неизвестна, однако она составляет не менее 1 кГц. На интересующей нас частоте отстройки $f' = 10$ кГц зависимость уровня фазовых шумов будет спадать пропорционально $1/f'^2$.

Слагаемые в формуле (3), содержащие граничную частоту фликкер-шума в дальнейших расчетах учитываться не будут. В формулу (3) подставим выражения (4), (5), а также известные параметры: $T = 300$ К, $N_f = 2,5$ дБ/Гц, $N_{RIN} = -145$ дБ/Гц. При величине отстройки $f' = 10$ кГц уровень шумов генератора при включении ВОТ1 (1000 м) составляет минус 136,4 дБ, а при включении ВОТ2 (30 м) всего минус 107,6 дБ. Однако при использовании более протяженных волоконно-оптических трактов уменьшается расстояние между соседними модами в спектре выходного сигнала.

Известно, что при включении двух волоконно-оптических трактов в схему генератора, результирующее расчетное значение нагруженной добротности будет определяться, как среднее арифметическое эквивалентных добротностей обоих трактов.

$$Q = \frac{4,6 \cdot 10^5 + 1,4 \cdot 10^4}{2} = 2,3 \cdot 10^5$$

Графики спектральной плотности мощности фазовых шумов одноконтурного при разных длинах ВОТ и двухконтурного генератора представлены на рис. 5.

На отстройке 10 кГц от центральной частоты генератора уровень фазовых шумов составляет минус 131,7 дБ. Этот результат на 5 дБ выше, чем у одноконтурного генератора при включении длинной волоконной линии. Однако при близких спектральных характеристиках двухконтурного генератора и одноконтурного при включении короткой волоконной линии двухконтурный имеет значительный выигрыш в плане селекции мод. Таким образом, двухконтурный оптоэлектронный генератор позволяет решить проблему компромисса между уровнем шумов и расстоянием между соседними модами генерации одноконтурных генераторов.

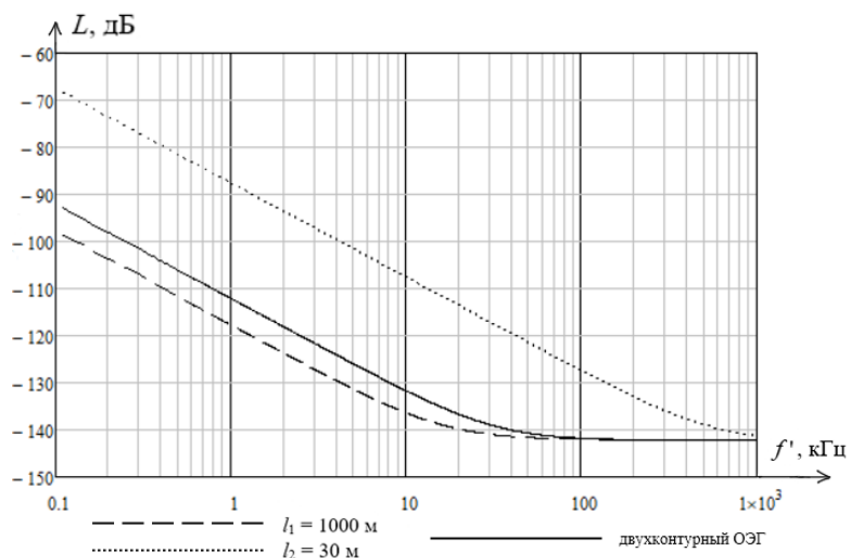


Рис. 5. СПМ фазовых шумов одноконтурного и двухконтурного генераторов

Заключение

Был создан макет двухконтурного ОЭГ с волоконно-оптическими трактами длиной 1000 м и 30 м. Имеющаяся элементная база не позволила уменьшить массогабаритные показатели генератора. Экспериментально изучены характеристики элементов схемы. Теоретически были оценены спектральные характеристики генератора, которые хорошо согласуются с результатами практического исследования. Также удалось рассчитать спектральную плотность мощности фазовых шумов одноконтурного и двухконтурного генератора, что подтвердило предположение о решении компромисса между уровнем шумов и требованием к полосе пропускания фильтра в одноконтурном оптоэлектронном генераторе.

Библиографический список

1. Белкин М.Е. Оптоэлектронный генератор СВЧ сигналов: моделирование, исследование спектральных и шумовых характеристик / М.Е. Белкин, А.В. Лопарев // Нано- и микросистемная техника. – 2011, – №9. – С. 29-33.
2. Ченаки А. Фазовые шумы в СВЧ-генераторах. Методы решения проблемы / А. Ченакин // ЭЛЕКТРОНИКА: Наука, Технология, Бизнес. – 2011, – №4. – С. 52-61
3. Leeson D. B. A simple model of the feedback oscillator noise spectrum / D.B. Leeson // Proceedings of the IEEE, 1966, Vol. 54, № 2. – pp.329–330.

**V.V. Biryukov, V.A. Grachev, S.A. Kapustin, Kulyaba D.G., M. S. Olkhova,
A.S. Raevskii, D.M. Shustov**

STUDY OF THE SPECTRAL CHARACTERISTICS OF A DOUBLE-LOOP OPTOELECTRONIC OSCILLATOR

Nizhny Novgorod state technical university n.a. R.E. Alekseev

In this article, we consider the realization of a double-loop optoelectronic oscillator of radio frequency signals; we study the level of the spectral power density of phase fluctuations, and the spectral characteristics of a single-loop and double-loop generator.

Keywords: optoelectronic generator, radio photonics, optical delay line, phase noise, autogenerator.

ИЗВЕСТНЫЕ И НОВЫЕ ПОДХОДЫ К ОПИСАНИЮ ШЕРОХОВАТЫХ ПОВЕРХНОСТЕЙ ВОЛНОВОДА

Нижегородский государственный технический университет им. Р.Е. Алексеева

В данной работе приводится описание моделей и методов, применяемых при описании шероховатых поверхностей. Наряду с достаточно известными методами приводятся подходы, предлагаемые авторами статьи.

Ключевые слова: шероховатость поверхности, экранирующие поверхности, волноводы, коэффициент затухания

Введение

На практике измеренные значения потерь в волноводной структуре превышают, рассчитанные с помощью известных теоретических формул, поскольку реальная структура стенок волновода имеет некоторую шероховатость, что увеличивает длину пути токов на стенках волновода, и, соответственно, поверхностное сопротивление проводника. В миллиметровом диапазоне толщина скин-слоя становится сопоставима со значением среднеквадратичного отклонения шероховатости материала волновода, что оказывает влияние на погонные потери. При повышении рабочей частоты требования к качеству обработки экранирующих поверхностей возрастают и становятся соизмеримыми с технологическими возможностями. В связи с этим, постоянно растёт интерес как к изучению влияния шероховатости поверхности на распространение волн в экранированных структурах с учетом конечной проводимости металла, так и к способам описания шероховатостей поверхности [1-3].

Известные способы описания шероховатых поверхностей проводников

Как правило, описание шероховатых поверхностей проводников производят, представляя профили шероховатостей в виде простых геометрических фигур. В некоторых моделях для упрощения анализа эти фигуры располагают периодически. В итоге оценка потерь из-за шероховатостей часто сводится к вычислению поправочных коэффициентов. Первой моделью для учета потерь, вызванных шероховатостью поверхности, стала модель Моргана, представляющая собой периодическую структуру типа зуба пилы [4]. Главное предположение в этой модели заключалось в том, что ток течет по краю шероховатой поверхности, что приводит к увеличению пути, проходимого током, а значит и к дополнительным потерям мощности. Морган, используя метод конечных разностей при решении задачи для своей структуры, вывел отношение потерь мощности, рассеянной в проводнике с шероховатостью, к мощности, рассеянной в проводнике с гладкой поверхностью. Несколько позже Хаммерстед и Джонсон получили эмпирическое выражение, основанное на результатах Моргана, используя только один параметр неровности — ее среднеквадратичную высоту [5]:

$$K = \frac{\alpha_{шер}}{\alpha_{глад}} = 1 + \frac{2}{\pi} \cdot \arctan \left(1.4 \left(\frac{h_{cp}}{\delta} \right)^2 \right),$$

Однако полученное выражение насыщается, когда значения коэффициента K достигает двойки. На практике же влияние шероховатости может быть значительно больше, что говорит о том, что этот метод применим далеко не всегда.

Иной способ описания шероховатой поверхности предложил Холм, его модель описывает неровности в виде полусфер, выступающих из плоскости проводника [6].

Задача рассеяния плоской волны от полусферического выступа на плоскости решается методом зеркальных отображений. В итоге поправочный коэффициент имеет вид:

$$K = \frac{\left| \operatorname{Re} \eta \left\{ \frac{3\pi}{4\kappa^2} (\alpha_1 + \beta_1) \right\} \right| + \frac{\mu_0 \omega \delta}{4} \left\{ d_{\text{пик}}^2 - \pi \left(\frac{b_{\text{основ}}}{2} \right)^2 \right\}}{\frac{\mu_0 \omega \delta}{4} d_{\text{пик}}^2}$$

где $d_{\text{пик}}$ - расстояние между пиками шероховатостей, $b_{\text{основ}}$ - длина основания неровности,

κ – волновое число, α_1 и β_1 - коэффициенты рассеяния первого порядка, $\eta = \sqrt{\frac{\mu_0}{\varepsilon_0 \varepsilon'}}$.

Данный метод основан на трех входных параметрах и требует трех статистических измерений. Более новая версия аналогичного подхода была предложена Хурэем. В его модели шероховатая поверхность моделируется в виде пирамидального набора сферических частиц на поверхности проводника [7,8]. Задача рассеяния и поглощения решается аналогично модели полусфер, но для каждой сферы отдельно. Обязательными параметрами являются диаметры сфер и их количество N , приходящееся на площадку. Используя принцип суперпозиции можно получить результирующие потери в подобной структуре. Сложность этого метода связана с тем, что поверхностные структуры, напоминающие это изображение если и можно наблюдать, то только на полученной электроосаждением медной фольге, а основные параметры модели - диаметр сфер внутри пирамидок и их количество на площадке - невозможно наблюдать при измерении.

Цанг с соавторами осуществил более сложный анализ проблемы шероховатости поверхности [9,10]. Они проанализировали двумерные случайные шероховатые поверхности на основе метода малых возмущений второго порядка и численного метода моментов, выполнили расчет коэффициента поглощения мощности для шероховатости поверхности с помощью гауссовой и экспоненциальной корреляционных функций. В результате они получили, что коэффициент, описывающий потери, зависит от трех параметров: среднеквадратичной высоты, длины корреляции и корреляционной функции. Позже этот подход был распространен на анализ шероховатости трехмерной поверхности, где высота поверхности изменяется в обоих горизонтальных направлениях.

Вероятно, самой свежей из интересных моделей для описания шероховатостей является градиентная [11]. Она не предполагает наличия резкой границы между диэлектриком и проводником, шероховатость моделируется как непрерывный переход проводимости от практически нуля в диэлектрике, до объемной проводимости металла. Поскольку профиль шероховатости для большинства поверхностей имеет нормальное распределение, градиент проводимости выражается, используя только один метрологический параметр шероховатости — среднеквадратичную высоту $h_{\text{ср}}$.

$$\sigma(x) = \sigma_{\text{мет}} \int_{-\infty}^x PDF du = \sigma_{\text{мет}} \int_{-\infty}^x \exp\left(\frac{-u^2}{2h_{\text{ср}}^2}\right) du,$$

Выражение для скин-эффекта на шероховатой поверхности выводят из уравнений Максвелла для гармонических полей. Предполагая проводник немагнитным с поверхностью в плоскости уз, можно прийти к дифференциальному уравнению только в одном измерении (B_y), в котором третье слагаемое получается из-за проводимости, зависящей от ее координаты $\sigma(x)$:

$$\frac{\partial^2 B_y}{\partial x^2} + j\omega\mu_0\sigma B_y - \frac{\partial}{\partial x} \ln \sigma \cdot \frac{\partial B_y}{\partial x} = 0.$$

Новые способы описания шероховатых поверхностей волноводов

Ранее в ряде работ, например, [1, 3, 12] был предложен подход, основанный на концепции парциальных волн Бриллюэна и позволяющий учесть случайную шероховатость поверхности произвольной формы, причем этот подход был расширен как на волноводы прямоугольной формы, так и круглые волноводы. В соответствии с данным представлением задача нахождения затухания собственных волн направляющих электродинамических структур с шероховатыми экранирующими поверхностями сводится к достаточно хорошо исследованному случаю структур с идеально гладкими поверхностями. Эквивалентность реальной шероховатой поверхности и гладкой поверхности достигается внесением поправки в удельную проводимость материала последней, обеспечивающей равенство коэффициентов отражения электромагнитной волны от рассматриваемых поверхностей. Средний коэффициент отражения от шероховатой поверхности и коэффициент отражения от гладкой поверхности связаны соотношением:

$$\bar{\Gamma}(\theta) = \Gamma(\theta) \int_{-\infty}^{\infty} \exp(ik_0 2\xi \cos\theta) \omega(\xi) d\xi$$

где Γ - коэффициент отражения от гладкой поверхности, θ – угол падения волны, ξ – высота шероховатости, $\omega(\xi)$ – плотность вероятности.

Тем самым найти затухание волны, вызванное шероховатостью поверхности, можно, внося поправки в удельную проводимость материала экрана.

Еще один из способов был описан в работе [13], где предлагается модификация, описанного ранее градиентного метода для определения влияния шероховатости поверхности на затухание волн в экранированных структурах. Любую шероховатую поверхность можно разбить на слои, приняв за начало разбиения линию, проходящую через точку максимальной глубина впадины профиля, и сделав последним слоем разбиения слой, включающий в себя точку максимальной высоты пика профиля. При этом эффективная проводимость каждого получившегося слоя будет находиться в интервале от проводимости металла до проводимости воздуха, изменяясь от первого слоя к N-ому. Характер изменения эффективной проводимости от слоя к слою определяется профилем распределения шероховатостей. Исследование отражения электромагнитной волны от рассматриваемой шероховатой поверхности проводится методом ориентированных графов, который позволяет учесть многократные отражения волны от границ каждого слоя. Элементы матриц, описывающих границы между соседними слоями, находятся в соответствии с формулами Френеля:

$$\Gamma_{12} = \frac{Z_m(\omega, \sigma) - Z_{m+1}(\omega, \sigma)}{Z_m(\omega, \sigma) + Z_{m+1}(\omega, \sigma)},$$

$$\Gamma_{21} = \frac{Z_{m+1}(\omega, \sigma) - Z_m(\omega, \sigma)}{Z_{m+1}(\omega, \sigma) + Z_m(\omega, \sigma)},$$

$$T_{12} = \frac{2Z_m(\omega, \sigma)}{Z_m(\omega, \sigma) + Z_{m+1}(\omega, \sigma)},$$

$$T_{21} = \frac{2Z_{m+1}(\omega, \sigma)}{Z_m(\omega, \sigma) + Z_{m+1}(\omega, \sigma)},$$

где m – номер слоя разбиения, Γ - коэффициент отражения, T - коэффициент прохождения.

В результате можно строить матрицы рассеяния и матрицы прохождения для каждого полученного при разбиении слоя. Важным свойством матриц передачи является то, что матрица каскадного соединения ряда элементов равна произведению матриц передачи этих элементов. Таким образом, результирующая матрица передачи равна

произведению матриц передачи отдельных слоёв. Соответствующая ей матрица рассеяния даёт коэффициент отражения волны от рассматриваемой структуры, что позволяет найти эквивалентную проводимость рассматриваемой шероховатой поверхности, то есть проводимость идеально гладкой поверхности, имеющей такой же коэффициент отражения, как и шероховатая поверхность. Такое определение эквивалентной проводимости, основанное на концепции парциальных волн Бриллюэна, кажется наиболее адекватным для экранированных направляющих структур. Так как в случае направляющих электродинамических структур материал экранирующих поверхностей имеет высокую удельную проводимость, то это позволяет аппроксимировать зависимость коэффициента отражения от частоты и получить приближённое аналитическое выражение для эквивалентной удельной проводимости, которую можно либо подставить в известные выражения для вычисления потерь в волноводах, либо использовать при моделировании направляющих структур в САПР. В случае основной волны прямоугольного волновода можно учесть шероховатость экранирующих поверхностей, подставив эквивалентную удельную проводимость в известное выражение для коэффициента затухания:

$$\alpha = \frac{\omega \sqrt{\frac{2}{\omega \sigma_{\text{экв}} \mu_0} \left[\frac{1}{2} + \frac{b}{a} \left(\frac{\pi c}{a \omega} \right)^2 \right]}}{cb \sqrt{1 - \left(\frac{\pi c}{a \omega} \right)^2}}$$

где a и b – поперечные размеры волновода.

С использованием данной методики проводился расчёт потерь в прямоугольном волноводе сечением $1,1 \times 0,55$ мм при шероховатостях экранирующих поверхностей 1 мкм, $0,5$ мкм и $0,1$ мкм для приведённых профилей шероховатостей. Результаты таких расчетов показали, что погонные потери волновода зависят не только от величины шероховатостей экранирующей поверхности, но и от их профиля.

Еще один подход заключается в рассмотрении среды распространения энергии с точки зрения атомистических представлений. Тогда любую среду можно рассматривать как вакуум, в который вкраплены частицы вещества. Под действием падающей волны и излучений соседних атомов внутри атомов возбуждаются колебания электронов и атомных ядер и, как следствие, атомы становятся источниками вторичных сферических волн. Их интерференция между собой и с падающей волной определяет все поле во всем пространстве.

В поле электромагнитной волны атомы приобретают меняющиеся во времени дипольные моменты. Пусть в вакууме вдоль оси X распространяется плоская монохроматическая волна $E_0 = e^{i(\omega t - k_0 x)}$, на пути которой встречается тонкий плоскопараллельный слой, состоящий из точечных неподвижных атомов. Дипольные моменты атомов слоя, которые возбуждаются падающей волной, можно представить в виде $p = p_0 e^{i(\omega t - k_0 \epsilon)}$, где ϵ – абсцисса слоя, а электрическое поле каждого слоя в некой точке впереди расположенной точке A будет: $\frac{k_0}{r} p_0 e^{i(\omega t - k_0 \epsilon - k_0 r)}$, где r определяет расстояние от диполя. Данное выражение нужно просуммировать по всем диполям слоя для определения результирующего электрического поля в точке A .

Если стенку волновода представить, как сумму таких плоскопараллельных слоев, расположения частиц в которых зависит от того относится этот слой к шероховатости или не относится, то можно оценить какая часть энергии передается при отражении волны от одной стенки волновода к другой. Характер изменения проводимости (то есть плотности распределения атомов) от слоя к слою определяется профилем распределения шероховатостей.

Закключение

В статье представлен обзор наиболее известных способов оценки потерь в электродинамических направляющих структурах с шероховатыми экранирующими поверхностями и новых, предложенных авторами данной статьи. Некоторые из описанных моделей применимы в современных системах автоматизированного проектирования микроволновых структур. Область применения каждого из методов различна и зависит от характера решаемой задачи, но многообразие подходов к описанию таких шероховатых структур лишний раз подчеркивает актуальность данной темы.

Библиографический список

1. Бирюков В.В., Грачев В.А., Лобин С.Г., “Круглый экранированный волновод с шероховатой внутренней поверхностью”, *Антенны*, 2018, №10, с. 54-59.
2. Kashurkin O., “Measurements and simulation of conductor-related loss of PCB transmission lines” Missouri University of Science and Technology, 2016.
3. Biryukov V., Grachev V., Karakozova I. E., Lobin S., Shcherbakov V., Estimation of losses per unit length in a rectangular waveguide with rough screening surfaces based on the concept of partial waves. ITM Web Conf., Vol. 30, 29th International Crimean Conference (CriMiCo'2019), 2019.
4. Morgan S. P., “Effect of surface roughness on eddy current losses at microwave frequencies”, *Journal of Applied Physics*, vol. 20, no. 4, pp. 352-362, Apr 1949.
5. Hammerstad E. and Jensen O., “Accurate models for microstrip computer-aided design,” in *IEEE MTT-S Int. Microw. Symp. Dig.*, Washington, DC, USA, May 1980, pp. 407–409.
6. Hall S., Pytel S., Huray P., Hua D., Moonshiram A., Brist G., and Sijercic E., “Multigigahertz causal transmission line modeling methodology using a 3-D hemispherical surface roughness approach,” *Microwave Theory and Techniques, IEEE Transactions on*, vol. 55, no. 12, pp. 2614-2624, Dec 2007.
7. Huray P. G., Oluwafemi O., Loyer J., Bogatin E., and Ye X., “Impact of copper surface texture on loss: A model that works,” *DesignCon 2010 Proceedings*, 2010.
8. Huray P. G. et al., “Fundamentals of a 3-D ‘snowball’ model for surface roughness power losses,” in *Proc. IEEE Workshop Signal Propag. Interconnects*, Genova, Italy, May 2007, pp. 121–124.
9. Tsang L., Gu X., and Braunisch H., “Effects of random rough surface on absorption by conductors at microwave frequencies,” *Microwave and Wireless Components Letters, IEEE*, vol. 16, no. 4, pp. 221-223, April 2006.
10. Gu X., Tsang L., Braunisch H., and Xu P., “Modeling absorption of rough interface between dielectric and conductive medium,” *Microw. Opt. Technol. Lett.*, vol. 49, pp. 7–13, Jan. 2007.
11. Gold G. and Helmreich K., “Surface impedance concept for modeling conductor roughness,” in *IEEE MTT-S Int. Microw. Symp. Dig.*, Phoenix, AZ, USA, May 2015, pp. 1–4.
12. Бирюков В.В. “Расчет потерь в прямоугольном волноводе с шероховатыми экранирующими поверхностями”, *Антенны*. 2016. №7. С. 53–57.
13. Бирюков В.В., Лобин С.Г., “Градиентный метод расчета потерь направляющей электродинамической структуры с неидеальными экранирующими поверхностями”, *Физика волновых процессов и радиотехнические системы*, 2020, Том 23, № 2, с. 14-20.

V.V. Biryukov, T.V. Kozhevnikova, S.G. Lobin

**KNOWN AND NEW APPROACHES TO THE DESCRIPTION OF ROUGH SURFACES
OF A WAVEGUIDE**

Nizhny Novgorod state technical university n.a.R.E. Alekseev

This paper describes the models and methods used to describe rough surfaces. Along with well-known methods, the approaches proposed by the authors of the article are presented.

Keywords: surface roughness, shielding surfaces, waveguides, attenuation coefficient

М.Н. Бисярин ¹, Т.В. Кожевникова ¹, В.А. Малахов ¹, А.Н. Панин ², Ю.В. Раевская ¹,
В.В. Щербаков ¹

РАСЧЕТ И ЭКСПЕРИМЕНТАЛЬНОЕ ИССЛЕДОВАНИЕ РУПОРНОЙ ПИРАМИДАЛЬНОЙ АНТЕННЫ НА 220 ГГц

¹Нижегородский государственный технический университет им. Р.Е. Алексеева

²Институт физики микроструктур РАН

Произведено моделирование рупорного излучателя с использованием САПР Ansoft HFSS. Обоснована необходимость учёта в модели реальной толщины стенок рупора. Исследовано влияние конструктивных особенностей рупорной пирамидальной антенны на ее характеристики направленности на частоте 220 ГГц. Произведена экспериментальная проверка результатов расчёта.

Ключевые слова: рупорная антенна, фазовый центр, фазовая характеристика антенны, диаграмма направленности, коэффициент усиления, дальняя зона.

Введение

Терагерцовый диапазон частот перспективен для использования для беспроводных систем коммуникаций. Данный диапазон, благодаря высокой несущей частоте, позволяет передавать большие объемы данных.

В качестве антенн для передающих и приемных устройств терагерцового диапазона могут использоваться рупорные антенны как в качестве самостоятельной антенны в беспроводной системе связи на короткие расстояния, так и в качестве облучателей в зеркальных антеннах – в системах связи на дальние расстояния.

Ввиду того, что генераторы терагерцового диапазона, как правило, имеют невысокую мощность, на антенны накладывается дополнительное требование - их высокая эффективность. У высокоэффективных антенн большая доля излученной мощности сосредоточена в главном лепестке диаграммы направленности (ДН). Коэффициент рассеяния таких антенн должен быть низким. Указанное требование приводит к усложнению конструкции антенны.

В данной работе обсуждаются вопросы расчета и измерения характеристик пирамидальных рупорных антенн разных конструкций.

Объекты исследований и методика эксперимента

В докладе рассматривается антенна в виде пирамидального рупора, работающая на частоте 220 ГГц. Размеры рупора приведены на рис. 1, длина рупора 20 мм.

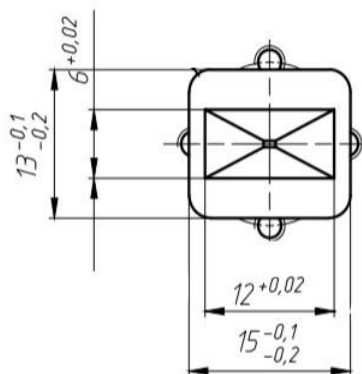


Рис. 1. Размеры пирамидальной рупорной антенны

Произведено сравнение результатов, получаемых для расчётных моделей с бесконечно тонкими стенками рупора и стенками конечной толщины (рис.1).

Расчет проводился с помощью САПР Ansoft HFSS. На рис. 2,3 приведены результаты расчёта амплитудных и фазовых характеристик рупорной антенны на частоте 220ГГц в плоскости E: для модели, учитывающей реальную толщину стенок, которая составляла 1,5мм (рис.2), и для упрощённой модели с бесконечно тонкими стенками (рис.3).

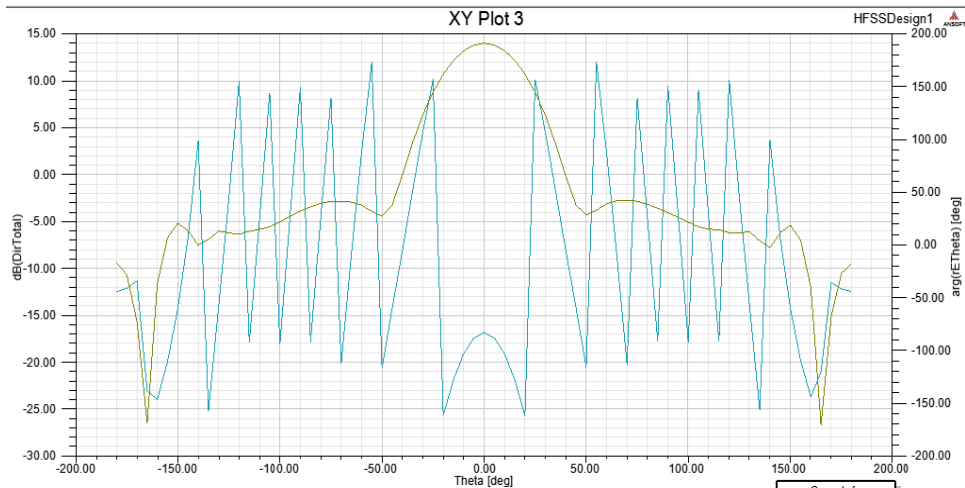


Рис. 2. Амплитудные и фазовые характеристики пирамидального рупора на частоте 220 ГГц с учетом толщины стенки

Из рис.2 видно, что при учёте реальной толщины стенок рассчитанная ДН антенны имеет главный лепесток, находящийся на «пьедестале», что говорит о низкой эффективности антенны. Очевидно, это является следствием антенно-фидерного эффекта, обусловленного влиянием токов, протекающих на торцевых поверхностях стенок рупора. Упрощенная модель, рис.3, данный эффект учесть не может, следовательно, является грубо приближённой. Для устранения растекания токов по торцевой поверхности рупора применяют систему из нескольких четверть волновых канавок, прорезаемых на торцах рупора.

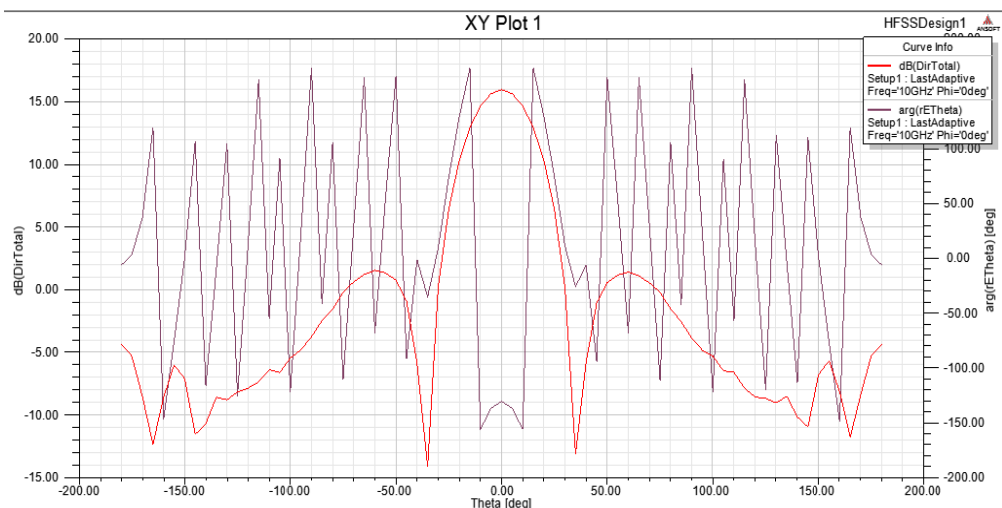


Рис. 3. Амплитудные и фазовые характеристики пирамидального рупора на частоте 220 ГГц без учета толщины стенок антенны

Исследуя фазовые характеристики антенн видно, что фазовый центр их явно невыражен. Но при этом, у антенны с толстыми стенками фазовый центр выражен хуже.

Для антенны с идеально тонкими стенками ширина главного лепестка ДН равна 11,62, для антенны с толстыми стенками 12,24 соответственно. Разница составляет чуть более 5 %.

Для проверки результатов расчета было выполнено экспериментальное исследование диаграммы направленности рупорной антенны с помощью методики измерения в дальней зоне. На рис. 4,а представлена упрощенная структурная схема измерительной установки [1]. Исследуемая антенна 3 (работающая в режиме приема) расположена на опорно-поворотном устройстве 4. На некотором расстоянии в дальней зоне этой антенны расположена вспомогательная передающая антенна 2, возбуждаемая генератором 1. Приемное устройство 5 имеет индикатор мощности 6, поступающей на вход приемного устройства. Зависимость показаний этого индикатора p от угла поворота антенны 4 при некотором фиксированном угле поворота φ_1 является сечением диаграммы направленности антенны по мощности $p(\theta, \varphi_1)$ в плоскости $\varphi_1 = \varphi_1 = \text{const}$. Выбирая различные значения угла $\varphi_1 = \varphi_1, \varphi_2, \dots$, можно измерить ДН в различных сечениях. При проведении измерений важно обеспечить отсутствие отражений от земли (см. штриховую линию на рис.4, а) и других окружающих предметов. Для ослабления влияния земли при измерениях антенны A_1, A_2 желательно размещать как можно выше над землей, а на земную поверхность укладывают щиты с поглощающим электромагнитное поле покрытием.

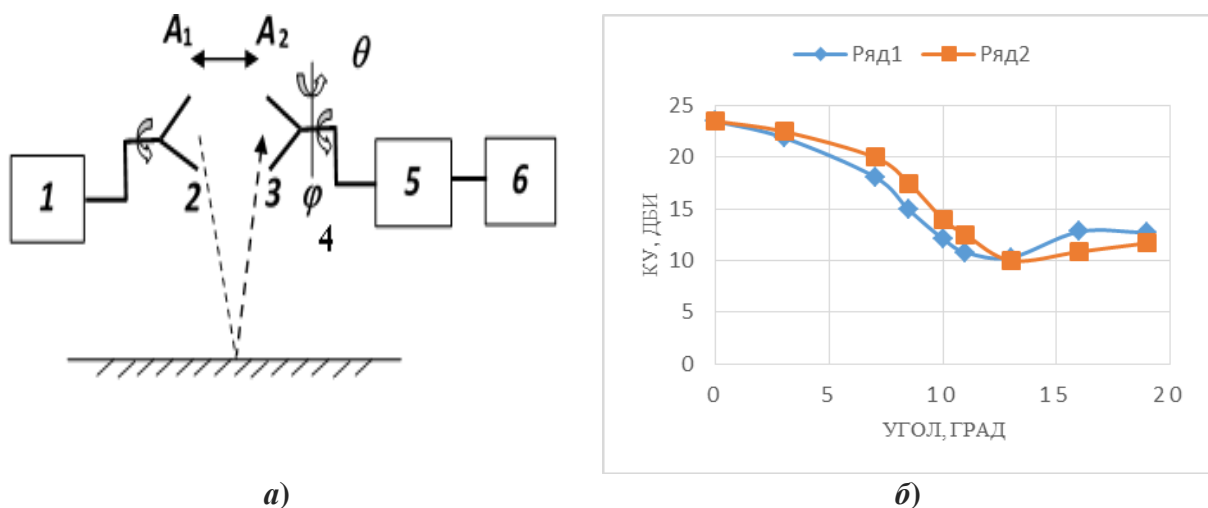


Рис. 4. Экспериментальное исследование антенны:

а - схема метода полигонных измерений; б – результаты измерений ДН в плоскости Е

На рис. 4,б приведена ДН пирамидальной рупорной антенны с апертурой 6x12 мм на частоте 220 ГГц в плоскости Е; ряд1 – измененная, ряд2 – расчетная. Максимальное отличие расчетных и экспериментальных значений не превышает 2,5дБ, что говорит о высокой эффективности используемого при расчетах САПР AnsoftHFSS.

Заключение

В данной работе приведены результаты расчета характеристик пирамидальной рупорной антенны на частоте 220 ГГц с помощью программного пакета САПР Ansoft HFSS. Обоснована необходимость использования расчётной модели, учитывающей реальную толщину стенок рупора. Произведено сравнение результатов расчета и эксперимента.

В ходе выполнения работы было произведено моделирование и расчет рупорных антенн с различной толщиной стенки с использованием САПР Ansys HFSS, т.к.

аналитически подобную задачу решить невозможно, ввиду особенной конструкции антенны, которые не учитываются при аналитическом решении.

Из результатов, полученных при моделировании видно, что характеристики антенны с толстыми стенками хуже, чем у антенны с идеально тонкими стенками. Коэффициент усиления антенны с толстыми стенками меньше на 1 %, чем у антенны с тонкими стенками. Ширина главного лепестка ДН у антенны с толстыми стенками 12.24° , что на 5% больше, чем у антенны с тонкими стенками. Эффективность рупорной антенны с широкими стенками хуже, чем с тонкими. Для ее улучшения необходимо усложнять конструкцию рупора, нанося четвертьволновые канавки на его торцевой поверхности.

Библиографический список

1. Методы измерений характеристик антенн СВЧ/ Под ред. Н.М. Цейтлина. М.: Радио и связь, 1985. – 368 с.

M.N. Bisyarin ¹, T.V. Kozhevnikova ¹, V.A. Malakhov ¹, A. N. Panin ², Yu.V. Raevskaya ¹,
V.V. Shcherbakov ¹

CALCULATION AND EXPERIMENTAL STUDY OF A PYRAMIDAL HORN ANTENNA AT 220 GHZ

¹*Nizhny Novgorod state technical university n.a.R.E. Alekseeva*

²*Institute for Physics of Microstructures, Russian Academy of Sciences*

Purpose: The influence of the design features of a pyramidal horn antenna on its directivity characteristics at a frequency of 220 GHz is investigated. An experimental check of the calculation results was carried out.

Design/methodology/approach: Simulation of a horn radiator was carried out using CAD Ansoft HFSS. An experimental study of the directional pattern of a horn antenna was carried out using the measurement technique in the far zone.

Findings: The performance of a thick-walled antenna is inferior to that of a perfectly thin-walled antenna. The gain of a thick-walled antenna is 1% less than that of a thin-walled antenna. The width of the main lobe of the pattern for a thick-walled antenna is 12.24° , which is 5% more than that for a thin-walled antenna.

Research limitations/implications: The performance of a horn antenna with wide walls is inferior to that of thin ones. To improve it, it is necessary to complicate the structure of the horn by applying quarter-wave grooves on its end surface.

Originality/value: This paper presents the results of calculating the characteristics of a pyramidal horn antenna at a frequency of 220 GHz using the Ansoft HFSS CAD software package. The necessity of using a computational model that takes into account the real thickness of the horn walls is substantiated. Comparison of calculation and experiment results is made.

Keywords: horn antenna, phase center, phase response of the antenna, directional pattern, gain, far zone.