

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



**приоритет2030<sup>+</sup>**  
лидерами становятся

# «ИНФОРМАЦИОННЫЕ СИСТЕМЫ И ТЕХНОЛОГИИ» ИСТ–2022

*Сборник материалов XXVIII Международной  
научно-технической конференции*

© Нижегородский государственный  
технический университет  
им. Р.Е. Алексеева, 2022

Нижний Новгород, 2022

**УДК 621:681**  
**ББК 32.97**  
**И650**

**Информационные системы и технологии - 2022:** [Электронный ресурс]: сборник трудов XXVIII Международной научно-технической конференции – Электрон. дан. – Н. Новгород: Нижегород. гос. техн. ун-т им. Р.Е. Алексеева, 2022. – 1 электрон. диск (CD-ROM): зв., цв., 12 см. – Систем. требования: ПК с процессором 486; ОЗУ 8 Мб.; операц/ сис-тема Windows 95; CD-ROM дисковод; мышь. – Загл. с экрана. – 200 экз.

В сборнике представлены материалы докладов XXVIII Международной научно-технической конференции ИСТ-2022, проведенной 23, 24 апреля 2022 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 12 городов РФ - преподавателей, научных сотрудников, докторантов, аспирантов, магистрантов и студентов - участников НИРС, а также сотрудников МГТУ им. Н.Э. Баумана, МЭИ, МГТУ «СТАНКИН», Высшей школы экономики, 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 20 предприятий и НИИ: ИПУ РАН, РФЯЦ-ВНИИЭФ, ИПФ 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, О.Г. Берестнева, Д.В. Жевнерчук, В.Ю. Карпычев, В.Р. Милов, Д.А. Ляхманов, С.Л. Моругин, А.С. Раевский, А.Г. Рындык, Э.С. Соколова, О.П. Тимофеева, Ю.М. Туляков, М.В. Ульянов, А.Д. Филинских, В.Л. Ягодкин

**ISBN 978-5-502-01612-4**

© Нижегородский государственный  
технический университет  
им. Р.Е. Алексеева, 2022



## СОДЕРЖАНИЕ

	страницы
СЕКЦИЯ 1. РАДИОТЕХНИЧЕСКИЕ СИСТЕМЫ И УСТРОЙСТВА .....	4
СЕКЦИЯ 2. ПРОЕКТИРОВАНИЕ РАДИОЭЛЕКТРОННЫХ УСТРОЙСТВ ВЧ И СВЧ ДИАПАЗОНОВ .....	85
СЕКЦИЯ 2. ТЕЛЕКОММУНИКАЦИОННЫЕ СИСТЕМЫ .....	137
СЕКЦИЯ 4. ИНФОРМАЦИОННЫЕ ТЕХНОЛОГИИ (автоматизация проектирования) .....	251
СЕКЦИЯ 5.1. ТЕХНИЧЕСКАЯ КИБЕРНЕТИКА (системы обработки информации).....	311
СЕКЦИЯ 5.2. ТЕХНИЧЕСКАЯ КИБЕРНЕТИКА (интеллектуальные системы управления).....	361
СЕКЦИЯ 5.3. ТЕХНИЧЕСКАЯ КИБЕРНЕТИКА (информационное моделирование когнитивных процессов).....	433
СЕКЦИЯ 6. МАТЕМАТИЧЕСКОЕ МОДЕЛИРОВАНИЕ ГЕОФИЗИЧЕСКИХ ПРОЦЕССОВ.....	461

**АЛГОРИТМ УСТРАНЕНИЯ НЕОДНОЗНАЧНОСТИ ОЦЕНКИ ВЗАИМНЫХ  
ВРЕМЕННЫХ ЗАДЕРЖЕК МНОЖЕСТВЕННЫХ ИСТОЧНИКОВ РАДИОИЗЛУЧЕНИЯ  
ШИРОКОПОЛОСНЫХ СИГНАЛОВ****И. В. Гринь**

ORCID: 0000-0002-3078-8738 e-mail: yelowt@mail.ru

**О. А. Морозов**

ORCID: 0000-0002-8411-2645 e-mail: oa\_morozov@nifti.unn.ru

**Н. А. Пинегина**

ORCID: 0000-0001-6633-3017 e-mail: natasha\_pinegina@mail.ru

Нижегородский государственный университет им. Н.И. Лобачевского

Актуальной задачей радиолокации в настоящее время является определение координат источника радиоизлучения методами пассивной пеленгации в реальном масштабе времени, которые для своей реализации требуют определения навигационных параметров (в частности, взаимных временных задержек распространения сигналов). В случае наличия нескольких источников радиоизлучения неизбежно появляется проблема соотнесения наборов навигационных параметров конкретному из источников. В данной работе предложен алгоритм устранения неоднозначности соотнесения наборов взаимных временных задержек, основанный на критерии согласованности временных задержек. Данный алгоритм может быть успешно применен к набору временных задержек, получаемых при анализе взаимной функции неопределенности принятых сигналов. В случае широкополосных сигналов, когда величина доплеровского смещения много меньше ширины спектра, в данной работе предлагается использовать алгоритм, основанный на построении модифицированной взаимной функции неопределенности и позволяющий компенсировать искажения спектра сигнала связанные с эффектом Доплера. В работе представлены полученные зависимости расчетных вероятностей правильно ассоциированных наборов временных задержек с источниками излучения от отношения сигнал/шум в исследуемых сигналах.

**Ключевые слова:** взаимная функция неопределенности, критерий согласованности временных задержек, цифровая фильтрация.

**ALGORITHM OF DISAMBIGUATING MUTUAL TIME DELAYS ESTIMATION FOR  
MULTIPLE WIDEBAND EMISSION SOURCES****I. V. Grin, O. A. Morozov, N. A. Pinegina**

National Research Lobachevsky State University of Nizhny Novgorod

*Nizhny Novgorod, Russia*

In this paper, an algorithm for disambiguating the association of mutual time delays based on the criterion of time delays congruence is proposed. In the case of multiple radio sources presence, the problem of associating the sets of navigation parameters to a particular source inevitably arises. An algorithm for disambiguating the association of mutual time delays based on the time delays congruence criterion can be successfully applied to a set of time delays obtained by mutual ambiguity function approach. For the successful use of the algorithm in the case of wideband signals, it is proposed to use an ambiguity function calculating modification based on the extraction of narrowband channels. The paper presents the obtained dependences of the calculated probabilities of correctly associated sets of time delays with radiation sources on the signal-to-noise ratio in the studied signals. This work is a starting point in the field of disambiguation arising from the determination of the navigation parameters of multiple radio sources, in the case of narrowband and wideband signals. The value of the entire article is the development and research of a method for applying the algorithm for disambiguating the association of mutual time delays sets to wideband signals.

**Key words:** mutual ambiguity function, time delays congruence criterion, digital filtering.

## Введение

Развитие спутниковых систем связи, связанное с увеличением спектральной полосы современных широкополосных сигналов и применением технологии разделения каналов, приводит к возникновению новых задач разработки алгоритмов обработки широкополосных сигналов в условиях априорной неопределенности принимаемых сигналов, а также низкого отношения сигнал/шум (ОСШ).

Наиболее актуальной из всех возникающих задач является задача определения координат источника радиоизлучения (ИРИ) методами пассивной пеленгации в реальном масштабе времени. Решение данной проблемы актуально во многих сферах, например, для работы систем спутниковой связи, при проектировании спутниковых поисково-спасательных систем и других прикладных задачах радиофизики, радиотехники и радионавигации. Для реализации алгоритмов определения координат ИРИ необходима информация о навигационных параметрах, в частности информация о взаимных временных задержках (ВВЗ) распространения сигналов [1].

Традиционным методом определения ВВЗ распространения сигналов при наличии шума в исследуемых каналах является метод максимального правдоподобия [2]. В условиях существенного влияния эффекта Доплера метод максимума правдоподобия обычно реализуется на основе алгоритмов построения и анализа взаимной функции неопределенности (ВФН). Для пары принимаемых сигналов  $s_1(t)$  и  $s_2(t)$  ВФН имеет следующий вид [3]:

$$A_{12}(\Delta t, \Delta f) = \int_{-\infty}^{+\infty} s_1(t) s_2^*(t + \Delta t) \exp(-j2\pi\Delta f t) dt. \quad (1)$$

В случае узкополосных сигналов ВВЗ и доплеровский сдвиг могут быть определены по положению главного максимума ВФН:

$$(\Delta t, \Delta f) = \arg \max_{\Delta t, \Delta f} |A(\Delta t, \Delta f)|. \quad (2)$$

Следует отметить, что одновременное позиционирование нескольких ИРИ приводит к возникновению неоднозначности при соотнесении каждого из навигационных параметров конкретному из источников. В связи с данной проблемой представляется актуальной реализация алгоритмов устранения данной неоднозначности. В данной работе предлагается алгоритм устранения неоднозначности при определении ВВЗ, основанный на построении и анализе ВФН, который позволяет уменьшить время на обработку и анализ принятых сигналов.

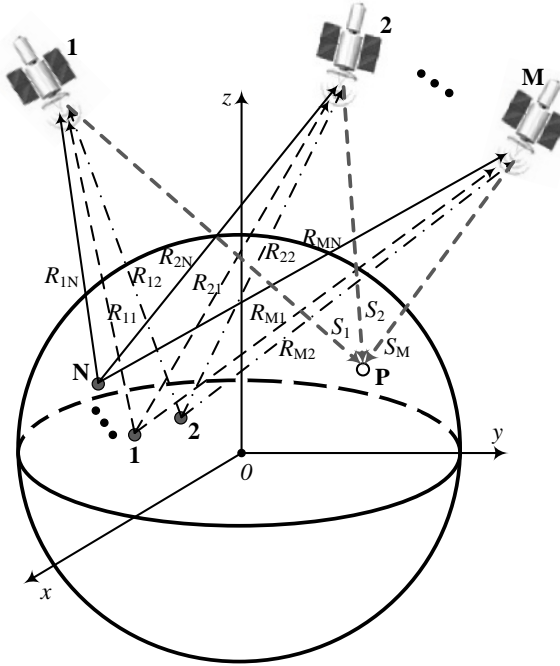
### Алгоритм устранения неоднозначности оценки временных задержек

На рис. 1 приведена многопозиционная спутниковая система пассивной пеленгации. Главная задача данной системы – определение координат  $(x_i, y_i, z_i)$   $N$  ИРИ, находящихся на поверхности Земли. Наиболее распространенным методом определения координат ИРИ является разностно - дальномерный метод, требующий для своей реализации определения навигационных параметров – ВВЗ распространения сигналов.

Модель сигналов, принимаемых спутниками, может быть представлена в следующем виде:

$$\left\{ \begin{array}{l} s_1(t) = \sum_{i=1}^N x_i(t) + \eta_1(t); \\ s_2(t) = \sum_{i=1}^N x_i((1 - \alpha_{2i})t + \tau_{2i}) + \eta_2(t); \\ s_j(t) = \sum_{i=1}^N x_i((1 - \alpha_{ji})t + \tau_{ji}) + \eta_j(t); \\ \dots \\ s_M(t) = \sum_{i=1}^N x_i((1 - \alpha_{Mi})t + \tau_{Mi}) + \eta_M(t). \end{array} \right. \quad (3)$$

где  $x_i(t)$ ,  $i=1, \dots, N$  – сигнал, излучаемый  $i$ -м источником, сигналы  $x_i((1-\alpha_{ji})t + \tau_{ji})$ ,  $j = 2, \dots, M$  представляют собой искажённые ( $\alpha_{ji} = V_r/c$  – коэффициент, определяющий влияние эффекта Доплера и равный отношению радиальной скорости спутника к скорости света) и задержанные копии сигнала  $x_i(t)$ ;  $\eta_j(t)$  – некоррелированные с сигналами аддитивные шумы в разных каналах распространения. Для определения координат источника радиоизлучения с помощью разностно-дальномерного метода необходима информация о временных задержках  $\tau_{ji}$  для каждого источника с номером  $i$ .



**Рис. 1. Схема многопозиционной спутниковой системы пассивной пеленгации**

При проведении оценки ВВЗ сигналов, принимаемых космическими аппаратами в описанной системе (3), на основе построения и анализа ВФН неизбежно возникает проблема неоднозначности соотнесения величин измеряемых задержек конкретным ИРИ [4].

При построении ВФН системы (3) каждая ВФН будет содержать  $N$  пиков.

$$\left\{ \begin{array}{l} A_{12}(\Delta t, \Delta f): \Delta t_{121}, \Delta t_{122}, \dots, \Delta t_{12N}; \\ A_{23}(\Delta t, \Delta f): \Delta t_{231}, \Delta t_{232}, \dots, \Delta t_{23N}; \\ \dots \\ A_{M-1,M}(\Delta t, \Delta f): \Delta t_{M-1M1}, \Delta t_{M-1M2}, \dots, \Delta t_{M-1MN}; \\ A_{M,1}(\Delta t, \Delta f): \Delta t_{M11}, \Delta t_{M12}, \dots, \Delta t_{M1N}. \end{array} \right. \quad (4)$$

Можно заметить, что рассмотренные задержки, соответствующие одному источнику, обладают следующим свойством:

$$\Delta t_{12k} + \Delta t_{23k} + \dots + \Delta t_{M1k} = \tau_{1k} - \tau_{2k} + \tau_{2k} - \tau_{3k} + \dots + \tau_{Mk} - \tau_{1k} = 0. \quad (5)$$

где  $\tau_{ij}$  – время распространения излучения от  $j$ -го источника до  $i$ -го космического аппарата.

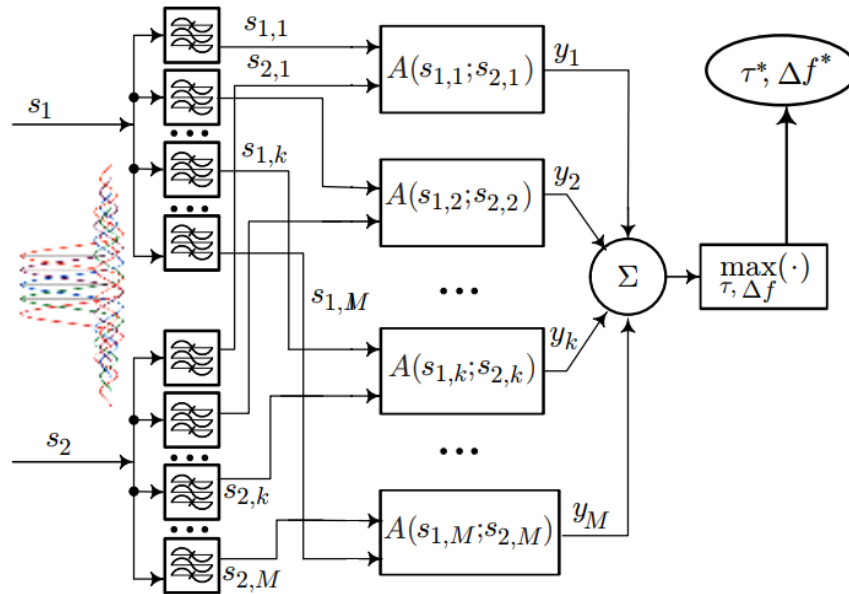
Таким образом, для устранения неоднозначности соотнесения величин измеряемых задержек конкретным источникам излучения, использующий критерий согласованности временных задержек (5), требуется оценить  $N$  ВВЗ и для каждой ВВЗ найти из остальных функций неопределенности системы (4) ВВЗ, удовлетворяющие критерию согласованности задержек (5).

В случае относительно широкополосных сигналов перед применением критерия согласованности временных задержек предлагается использовать алгоритм, основанный на

разбиении широкополосных сигналов  $s_1(t)$  и  $s_2(t)$  на  $M$  частотных каналов, например, используя алгоритм цифровой фильтрации [5]:

$$s_1^i[n] = \sum_{k=0}^L s_1[k]h_i[n-k], \quad s_2^i[n] = \sum_{k=0}^L s_2[k]h_i[n-k]. \quad (6)$$

Синтез цифровых фильтров  $\{h_i\}$ , настроенных на каждую из возможных центральных частот, может быть произведен заранее. После выполнения цифровой фильтрации набором фильтров, получим набор из  $M$  узкополосных сигналов. Структурная схема предлагаемого алгоритма определения взаимной временной задержки широкополосных сигналов представлена на рис. 2:



**Рис. 2. Схема алгоритма оценки взаимной временной задержки широкополосных сигналов**

Применяя данный алгоритм циклически для каждого номера  $i=1\dots N$ , можно определить взаимные временные задержки, соответствующие всем источникам в системе. Однако, в случае наличия нескольких ИРИ выраженность главных максимумов ВФН мала, что не позволяет производить качественный анализ. В данной работе предлагается модифицированная формула для построения ВФН, позволяющая компенсировать как смещение, так и масштабирование спектра широкополосного сигнала. Поскольку данные эффекты связаны с влиянием эффекта Доплера, возможно использовать разницу доплеровских коэффициентов ( $\Delta k$ ) для описания искажения спектров сигналов при построении ВФН:

$$A_2(\Delta t, \Delta k) = \int_{-\infty}^{+\infty} s_1(t\Delta k)s_2^*((t + \Delta t)\Delta k) \exp(-j2\pi f_0\Delta kt) dt. \quad (7)$$

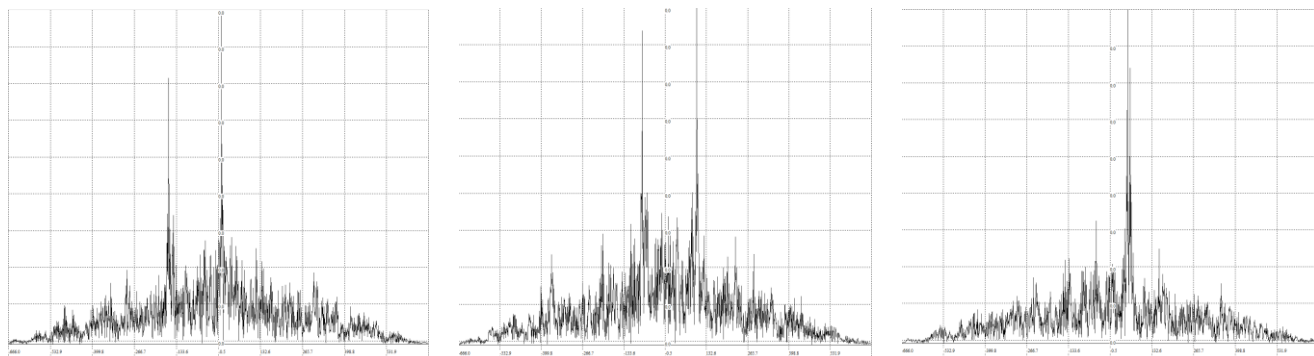
Дальнейшее устранение возникающей неоднозначности основывается на алгоритме, использующем критерий согласованности временных задержек, описанном ранее.

### Результаты моделирования

Для исследования предложенного алгоритма было проведено моделирование сигналов с псевдослучайной перестройкой рабочей частоты. Несущая частота сигналов изменялась случайным образом в полосе шириной 250 МГц (от 950 до 1200 МГц), разделенной на 51 частотный канал. Частота дискретизации сигналов принята равной 300 МГц, скорость передачи информации 10 Мбит/с. На основе компьютерного моделирования проведено исследование зависимости вероятности корректного соотнесения набора временных задержек источникам излучения от уровня шума (ОСШ) в исследуемых сигналах. Вероятность рассчитывалась на основе:  $P=m/n$ ,  $m$  – количество корректных срабатываний алгоритма,  $n$  – количество испытаний.

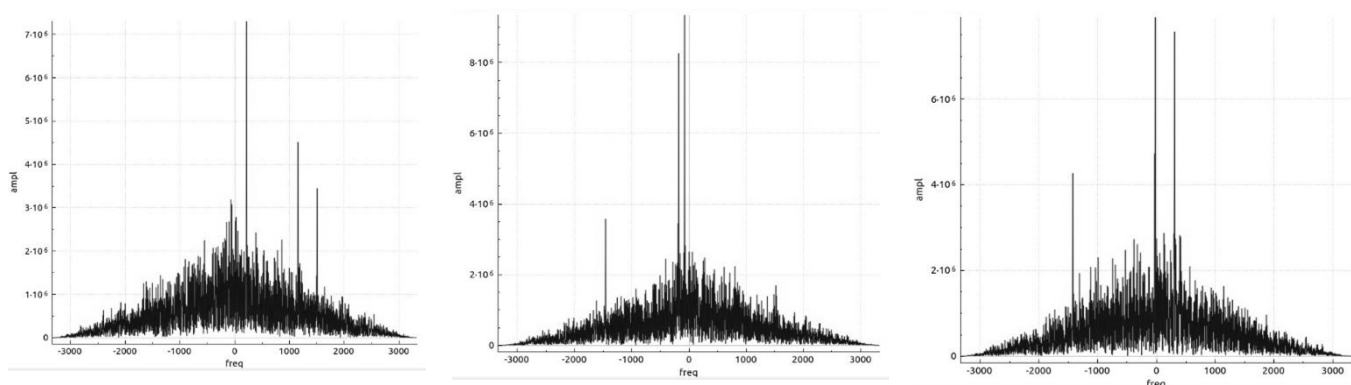
Испытание считалось благоприятным, если временные задержки определились верно, с точностью до половины длительности бита, а также учитывалось найденное количество взаимных временных задержек в каждой выделенной группе (количество взаимных временных задержек в каждой группе должно быть равно количеству источников). Усреднение проводилось по  $N = 1000$  реализациям. Исследовались конфигурации, содержащие 3 источника излучения.

На рис. 3 представлены «сечения» ВФН, полученных с помощью алгоритма, структура которого изображена на рис. 3 при наличии 3 ИРИ и 3 космических аппаратов.



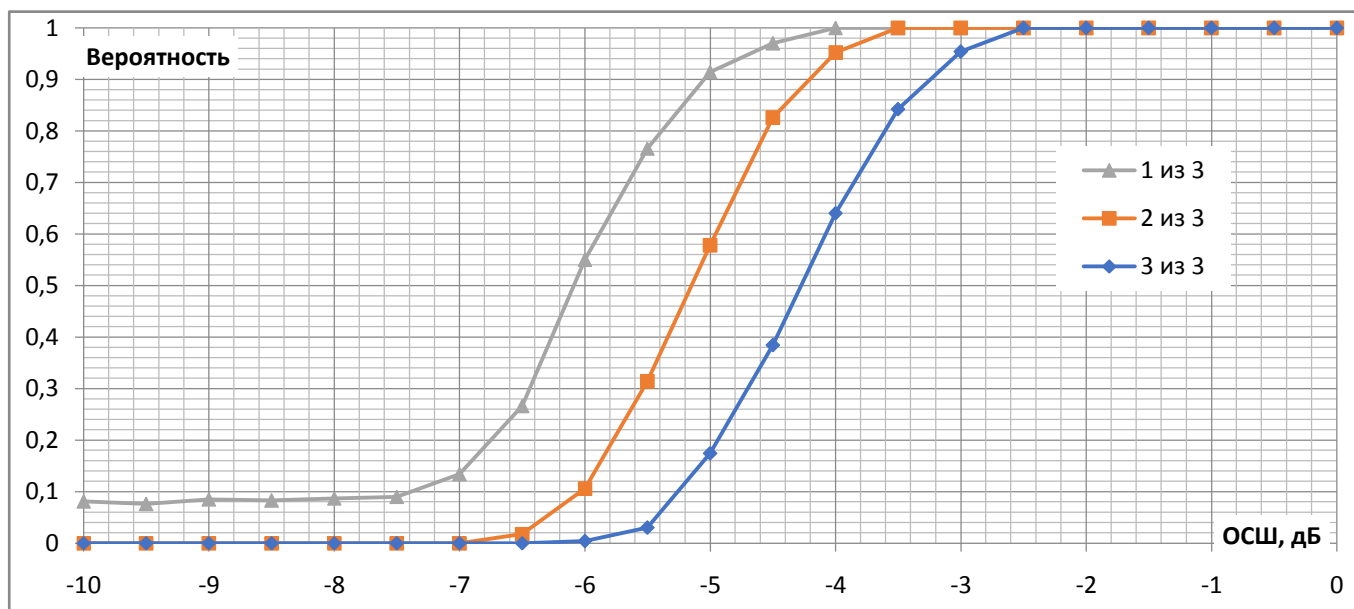
**Рис. 3. Вид «сечений» ВФН при ОСШ=+3дБ широкополосных сигналов с ППРЧ**

На рис. 4 представлены сечения ВФН, построенной на основе выражения (7), при наличии 3 ИРИ и 3 космических аппаратов имеют следующий вид:



**Рис. 4. Вид «сечений» модифицированной ВФН при ОСШ=+3дБ широкополосных сигналов с ППРЧ**

На основе компьютерного моделирования проведено исследование зависимости вероятности корректного соотнесения набора временных задержек источникам излучения от уровня шума (ОСШ) в исследуемых сигналах. Вероятность рассчитывалась на основе:  $P=m/n$ ,  $m$  – количество корректных срабатываний алгоритма,  $n$  – количество испытаний. Испытание считалось благоприятным, если временные задержки определены верно, с точностью до половины длительности бита, а также учитывалось найденное количество взаимных временных задержек в каждой выделенной группе (количество взаимных временных задержек в каждой группе должно быть равно количеству источников). Усреднение проводилось по 1000 реализаций. Исследовались конфигурации, содержащие 3 источника излучения. На рис. 5 представлены полученные зависимости. Приведены зависимости корректного соотнесения 1 из 3, 2 из 3 и 3 из 3 наборов ВВЗ источникам излучения.



**Рис. 5.** Зависимости вероятностей корректного соотнесения набора временных задержек источникам излучения от отношения сигнал-шум в исследуемых широкополосных сигналах

### Заключение

В данной работе предложен алгоритм устранения неоднозначности соотнесения наборов взаимных временных задержек, основанный на критерии согласованности временных задержек. Данный алгоритм может быть успешно применен к набору временных задержек, получаемых при анализе взаимной функции неопределенности принятых сигналов. В случае широкополосных сигналов, когда величина доплеровского смещения много меньше ширины спектра, в данной работе предлагается использовать алгоритм, основанный на построении модифицированной взаимной функции неопределенности и позволяющий компенсировать искажения спектра сигнала связанные с эффектом Доплера. В работе представлены полученные зависимости расчетных вероятностей правильно ассоциированных наборов временных задержек с источниками излучения от отношения сигнал/шум в исследуемых сигналах показывающие достаточную эффективность решения задачи соотнесения наборов навигационных параметров множественным источникам радиоизлучения.

### Библиографический список

1. **Гришин, Ю.П., Ипатов, В.П., Казаринов, Ю.М.** Радиотехнические системы: учеб. для вузов по спец. «Радиотехника»/ Под ред. Ю. М. Казаринова. – М.: Высш. шк., 1990. 496 с.
2. **Дженкинс, Г.Д., Ватте, Д.** Спектральный анализ и его приложения, Т.1. – М.: Мир, 1971, 319 с.
3. **Оппенгейм Э.** Применение цифровой обработки сигналов. – М.: Мир, 1980, – 553 с.
4. **Гринь, И.В., Ершов, Р.А., Морозов, О.А.** // Системы управления и информационные технологии, № 2.1 (56), 2014. – с.124.
5. **Гринь, И.В., Ершов, Р. А., Морозов, О. А.** // Изв. вузов. Радиофизика. 2019. Т. 62, № 10. С. 779.

## ПРИМЕНЕНИЕ ВИНЕРОВСКОЙ ФИЛЬТРАЦИИ В ЗАДАЧЕ ОПРЕДЕЛЕНИЯ ВЗАИМНОЙ ВРЕМЕННОЙ ЗАДЕРЖКИ СИГНАЛОВ

А.И. Зворыкин

ORCID: 0000-0002-8961-5962 e-mail: zvorykinnov@gmail.com

О.А. Морозов

ORCID: 0000-0002-8411-2645 e-mail: oa\_morozov@nifti.unn.ru

Нижегородский государственный университет им. Н.И. Лобачевского

В работе рассмотрена модификация метода оценки взаимной временной задержки сигналов с фазовой модуляцией на основе применения алгоритма оптимальной линейной фильтрации по критерию минимума среднеквадратической ошибки (винеровской фильтрации). В качестве заданного рассматривался сигнал канала дальности системы спутниковой связи с временным разделением доступа. Проведено компьютерное моделирование работы алгоритма обнаружения заданного сигнала в условиях влияния белого гауссова шума и райсовского шума. Получены зависимости правильного определения взаимной временной задержки сигналов от отношения сигнал/шум и сравнение результатов применения алгоритма на основе винеровской фильтрации с традиционным алгоритмом корреляционного обнаружения сигналов.

**Ключевые слова:** временная задержка сигналов, обнаружение сигналов, винеровская фильтрация, райсовский канал.

## APPLICATION OF WIENER FILTRATION IN THE PROBLEM OF DETERMINING SIGNALS WITH A MUTUAL TIME DELAY

A.I. Zvorykin, O.A. Morozov

National research Lobachevsky state university of Nizhny Novgorod

**Purpose:** This article describes a comparative analysis of the results of the phase modulation signal detection algorithms in the problem of determining the mutual time delay. Algorithms based on classical correlation estimation and optimal linear Wiener filtering were considered as detection methods.

**Design/methodology/approach:** To evaluate the effectiveness of the algorithms was used a computer simulation method. The result of the simulation is the dependence of the results of applying the calculations of the mutual delay of signals from the signal / noise ratio and comparing the results of detection algorithms. The simulation was carried out under the influence of white Gaussian noise and Rician noise on the receiving channel. As a given signal was used the signal of ranging from a satellite communication system with time division access.

**Findings:** The dependences of the probability of correct detection of a phase-modulated signal in the presence of various types of noise are obtained, a comparative analysis of the effectiveness of signal detection algorithms for various signal-to-noise ratios is carried out.

**Originality/value:** The results obtained in the course of the work make it possible to evaluate the effectiveness of using Wiener filtering in the problems of detecting signals with phase modulation against the background of various types of noise under conditions of a low signal-to-noise ratio. The proposed algorithm is further proposed to be used in conjunction with the nonlinear filtering method to compensate for the influence of the Doppler effect in the processing of signals from satellite communication systems. The results obtained are relevant for the development of signal detection algorithms in modern systems.

**Keywords:** signal time delay, signal detection, Wiener filtering, Rice channel



## Введение

Эффективность функционирования современных систем спутниковой связи напрямую связана с задачей определения местоположения источника радиоизлучения в условиях высокого уровня шума и доплеровского смещения частоты. Для решения данной задачи могут быть использованы различные подходы. Одним из наиболее часто применяемых методов является разностно-дальномерный метод [1], требующий оценки взаимных временных задержек принимаемых сигналов, распространяющихся по разным каналам в условиях низкого отношения сигнал/шум и доплеровского смещения частоты. Данная задача актуальна, например, для осуществления коммутации и установления сеанса связи космическими аппаратами спутниковых систем связи.

Традиционные алгоритмы решения подобных задач основаны на свойствах взаимной корреляционной функции принимаемых сигналов и на анализе спектра, что требует априорных сведений о спектральном составе сигналов и шумов. В общем случае задача определения взаимных временных задержек распространения сигналов при наличии эффекта Доплера, возникающего при распространении сигнала от движущегося источника, решается обобщенным методом максимального правдоподобия, сводящимся к вычислению и анализу взаимной функции неопределённости опорного  $s_1(t)$  и исследуемого  $s_2(t)$  сигналов [1]. В случае, когда смещением несущей частоты сигнала можно пренебречь, оптимальным методом обработки считается обобщенный коррелятор, при этом взаимная временная задержка определяется по положению максимума взаимной корреляционной функции сигналов. Эффективные алгоритмы обработки сигналов могут быть построены на основе применения различных методов цифровой фильтрации.

В работе рассмотрено применение корреляционной обработки и винеровской фильтрации [2] в задаче обнаружения заданных (ключевых) последовательностей фазоманипулированных сигналов и определения их взаимной временной задержки при различном уровне шумов для шумовых процессов различной природы.

### Алгоритмы обнаружения сигнала

В качестве наблюдаемого рассматривается сигнал вида

$$x(t) = \lambda \cdot s(t) + n(t), \quad (1)$$

где  $n(t)$  – помеховый стационарный гауссовский процесс с несмещённым средним и известной дисперсией,  $\lambda$  – коэффициент, характеризующий наличие или отсутствие полезного сигнала. Корреляционный принцип обнаружения для сигнала (1) может быть получен при вычислении логарифма отношения функции правдоподобия [1] и в отсутствие доплеровского смещения частоты сводится к вычислению взаимного корреляционного интеграла между наблюдаемым процессом (сигналом)  $x(t)$  и копией известного сигнала  $s(t)$

$$R_{xs} = \int_0^T x(t) \cdot s(t) dt. \quad (2)$$

Принцип оптимальной винеровской фильтрации заключается в поиске некоторого линейного фильтра  $h_g(\tau)$ , воздействие которого на входную выборку формирует сигнал с наименьшим отклонением от известного сигнала  $s(t)$  [2, 3]. Если в качестве критерия оценки отклонения используется средний квадрат ошибки

$$D = \left\langle \left[ s(t) - \int_0^T h_g(\tau) \cdot x(t - \tau) d\tau \right]^2 \right\rangle, \quad (3)$$

то при известных корреляционных функциях  $R_x$  и  $R_{xs}$  возможно произвести минимизацию функционала  $D$  вариационным принципом и, переходя в частотную область, определить передаточную характеристику оптимального винеровского фильтра [2]

$$H_e(f) = \frac{P_s(f) \cdot H^*(f)}{|H(f)|^2 \cdot P_s(f) + P_n(f)} = \frac{H^*(f)}{|H(f)|^2 + \frac{P_n(f)}{P_s(f)}}, \quad (4)$$

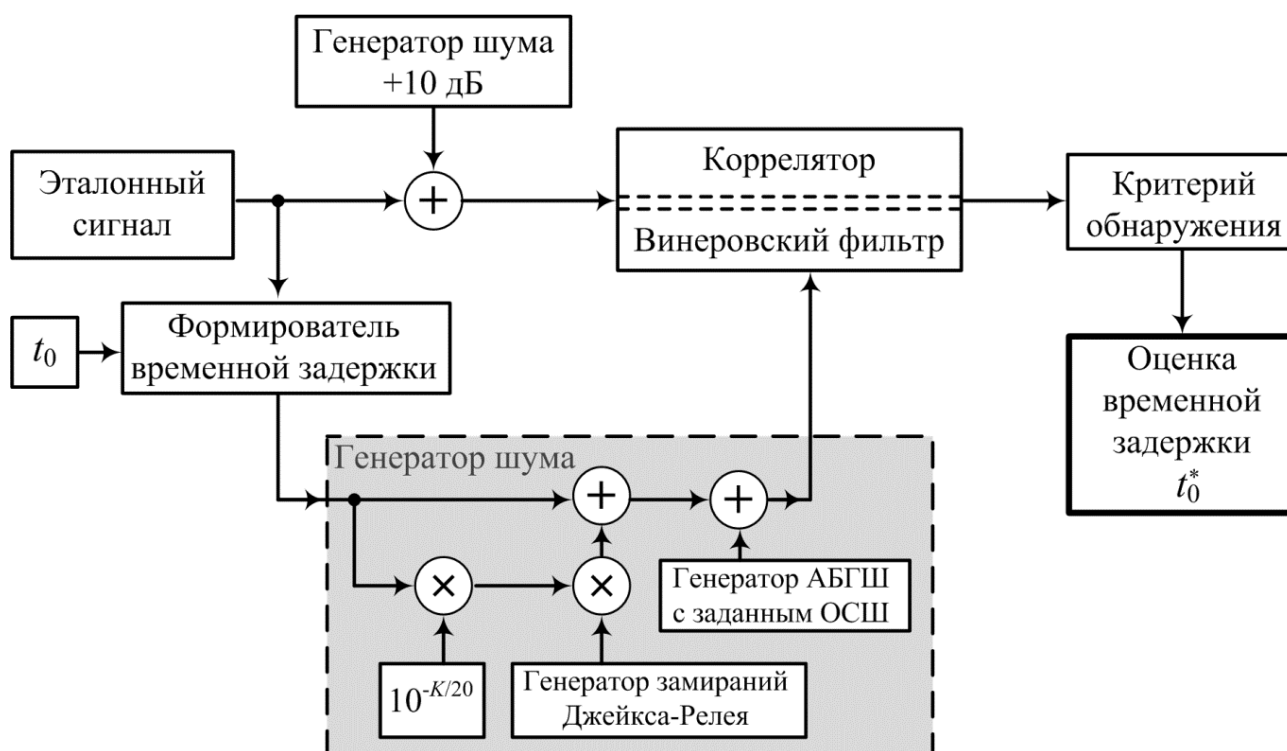
где  $H(f)$  – преобразование Фурье сигнала  $s(t)$ ,  $P_s(f)$  и  $P_n(f)$  – спектральные плотности мощности сигнала и шума.

При использовании модели аддитивного белого гауссова шума (АБГШ) выражение (4) может быть упрощено [3]. Выход оптимального линейного фильтра для полученной передаточной характеристики может быть представлен в виде

$$W(t) = \mathfrak{Z}^{-1}[H_e(f) \cdot X(f)], \quad (5)$$

где символом  $\mathfrak{Z}^{-1}$  обозначена операция обратного преобразования Фурье.

Наряду с моделью АБГШ для описания помех в современных системах спутниковой связи и мобильной цифровой связи используется модель однолучевого райсовского канала с замираниями [4, 5]. Наличие в канале райсовского шума можно представить в виде блок-схемы, рекомендованной стандартом GMR-1 [5]. Структурная схема обработки сигнала с учетом райсовского канала представлена на рис. 1.



**Рис. 1. Структурная схема алгоритмов обработки сигнала с использованием моделей АБГШ и райсовского канала**

Таким образом, для учета многолучевого распространения сигнала в системе спутниковой связи из входного комплексного сигнала (траектория распространения сигнала по прямой видимости) формируется дополнительный сигнал с заданными параметрами. Величина  $K$  на рис. 1 определяет отношение мощности прямого сигнала к общей мощности многолучевого распространения.

### Результаты моделирования

Для проведения сравнительного анализа обнаружителей, основанных на корреляционном принципе и оптимальной винеровской фильтрации, проведено исследование помехоустойчивости алгоритмов обнаружения и определения взаимной

временной задержки сигналов на основе имитационного компьютерного моделирования. В качестве несущего информационного сигнала  $s_1(t)$  был выбран сигнал канала дальности системы спутниковой связи с временным разделением доступа [6]. Используемый тип модуляции – фазовая модуляция ФМ2, скорость передачи данных внутри канала 9600 бит/с. Критериями, на основании которых принимается решение о наличии полезного сигнала в наблюдаемой временной выборке, выбраны превышения порогов для максимумов функций  $R_{xs}$  для корреляционного и  $W$  для оптимального винеровского обнаружителей.

Эталонный сигнал представляет собой короткую битовую последовательность, определяемую информационной составляющей канала дальности, в виде  $IQ$  – компонент. Используемая при моделировании битовая последовательность содержит 74 бита кодового слова и 6 бит информационного слова, задаваемого псевдослучайной последовательностью. Спектральные плотности мощности  $P_s(f)$  и  $P_n(f)$  в выражении (4) для винеровского фильтра определялись путем статистического моделирования сигнального и шумового (для райсовского канала) процессов.

Исследуемый сигнал  $s_2(t)$  представляет собой сдвинутую во времени на заданную величину задержки  $t_0$  копию эталонного сигнала, помещенную внутри псевдослучайной последовательности бит с такой же скоростью модуляции:

$$s_1(t) = x(t) + \xi(t),$$

$$s_2(t) = x(t - t_0) + \eta(t),$$

где  $\xi(t)$  и  $\eta(t)$  – некоррелированный с сигналом белый гауссов шум или райсовский шум с заданным отношением сигнал/шум (ОСШ).

Процесс моделирования задачи обнаружения и определения временной задержки сигнала, сформированного по битовой последовательности канала дальности, можно представить следующим образом.

1. Формирование временной задержки  $t_0$  путём встраивания сгенерированного известного сигнала в псевдослучайную битовую последовательность в заданный момент времени.

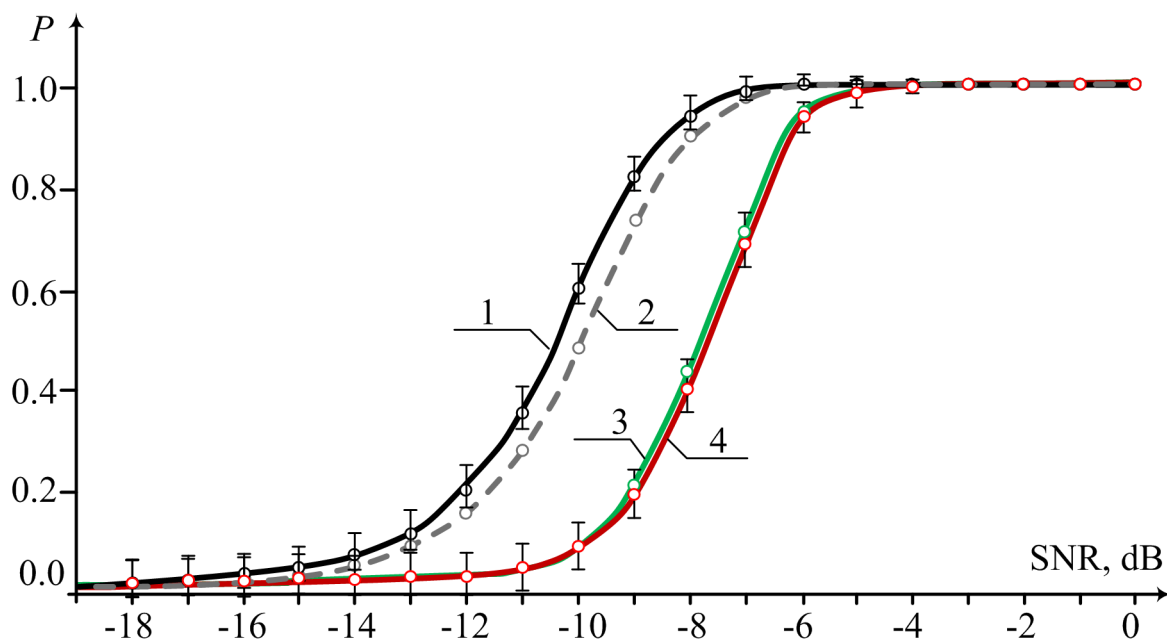
2. Наложение на полученный в п. 1 сигнал помехи, представляющей собой аддитивный белый гауссов либо райсовский шум с заданным ОСШ. Отношение сигнал/шум в полезном сигнале составляет фиксированную величину +10 дБ.

3. Вычисление взаимной корреляционной функции и выхода оптимального винеровского фильтра.

4. Определение по критерию обнаружения момента времени, соответствующего началу реализации полезного сигнала.

5. Оценка достоверности временной задержки сигнала – задержка считается правильно определённой, если значение измеренной задержки попадает в доверительный интервал, равный длительности одного символа ФМ-сигнала.

Зависимость вероятности правильного обнаружения и определения взаимной временной задержки сигнала  $P$  получена путем многократного запуска процесса обнаружения при различных фиксированных значениях отношения сигнал/шум в исследуемом сигнале в диапазоне от –20 дБ до 0 дБ. Графики вероятности правильного определения взаимной временной задержки сигнала для различных шумовых процессов приведен на рис. 2.



**Рис. 2. Зависимость вероятности обнаружения сигнала от ОСШ для длительности эталонного сигнала 80 бит при различных видах помех**

Линия 1 соответствует графику обнаружения оптимальной винеровской фильтрацией сигнала с аддитивным белым гауссовым шумом, линия 2 – графику обнаружения корреляционным алгоритмом сигнала с аддитивным белым гауссовым шумом. Линия 3 соответствует графику обнаружения оптимальной винеровской фильтрацией сигнала в райсовском канале, линия 4 – графику обнаружения корреляционным алгоритмом сигнала в райсовском канале.

Вероятность правильного определения взаимной временной задержки в условиях воздействия АБГШ близка к 1 при ОСШ около  $-8$  дБ. При наличии в канале райсовских замираний сигнала характеристики помехоустойчивости обоих обнаружителей ухудшаются примерно на  $2 \dots 3$  дБ.

### Выводы

В работе рассмотрена модификация метода оценки взаимной временной задержки сигналов с фазовой модуляцией на основе применения алгоритма оптимальной линейной фильтрации (винеровской фильтрации). На основе данного алгоритма решена задача определения взаимной временной задержки в условиях, когда доплеровским смещением спектра одного из обрабатываемых сигналов можно пренебречь. На основе компьютерного моделирования получены зависимости правильного определения взаимной временной задержки сигналов от отношения сигнал/шум и сравнение результатов применения алгоритма винеровской фильтрации с традиционным алгоритмом корреляционного обнаружения сигналов.

Анализ полученных зависимостей свидетельствует о более эффективной работе оптимального винеровского обнаружителя в области низких значений ОСШ (помехоустойчивость выше примерно на  $2$  дБ). Однако наличие в канале райсовских замираний сигнала приводит к ухудшению характеристик обоих обнаружителей.

В дальнейшем предполагается исследовать возможность применения винеровской фильтрации в задаче определения взаимных временных задержек узкополосных сигналов с учетом доплеровского смещения частоты. Для компенсации смещения спектра исследуемого сигнала предполагается использовать различные алгоритмы цифровой предварительной обработки сигналов [7, 8]. Цифровая фильтрация может быть использована для предварительной обработки фазоманипулированных сигналов с целью выделения моментов

манипуляций, соответствующих передаваемому коду. Применение данного подхода позволит значительно повысить вычислительную эффективность алгоритма определения взаимных временных задержек сигналов спутниковых систем связи в условиях влияния эффекта Доплера по сравнению с алгоритмом вычисления взаимной функции неопределенности.

#### **Библиографический список**

1. **Гришин, Ю.П., Казаринов, Ю.М., Ипатов, П.В.** Радиотехнические системы: Учеб. для вузов по спец. «Радиотехника» – М.: Высш.шк., 1990. – 496 с.
2. **Айфичер, Э., Джервис, Б.** Цифровая обработка сигналов. Практический подход. – Москва: Вильямс, 2004. – 992 с.
3. **Гельгор, А.Л., Горлов, А.И., Попов, Е.А.** Общая теория связи. Проверка статистических гипотез. Оценивание параметров. Оптимальный приём сигналов. – СПб.: Изд-во Политехн. ун-та, 2013. – 227 с.
4. **Ермолаев, В.Т., Флакман, А.Г.** Теоретические основы обработки сигналов в беспроводных системах связи: Монография. – Нижний Новгород: Изд-во ННГУ им. Н.И. Лобачевского, 2011. – 368 с.
5. ETSI TS 101 376-5-5 V1.3.1 (2005-02) GEO-Mobile Radio Interface Specifications (Release 1); Part 5: Radio interface physical layer specifications; Sub-part 5: Radio Transmission and Reception; GMR-1 05.005.
6. **Hitesh, A. Momaya, Varun, M. Patel, Vijay, I. Patel, Vijay, B. Patel.** Demand Assigned Multiple Access Subsystem // International Journal for Innovative Research in Science and Technology. – 2016. – Vol. 2, Issue 9 – P. 190-193.
7. **Логинов, А.А., Морозов, О.А., Солдатов, Е.А., Хмелев, С.Л.** Комбинированная цифровая фильтрация гармонического заполнения фазоманипулированных сигналов в задаче определения временной задержки // Известия высших учебных заведений. Радиофизика. – 2007. Т. 50. № 3. – С. 255-264.
8. **Логинов, А.А., Морозов, О.А., Хмелев, С.Л.** Алгоритм цифровой предварительной обработки сигналов с квадратурной фазовой модуляцией в задаче определения взаимной временной задержки // Известия высших учебных заведений. Радиофизика. – 2009. Т. 52. № 5-6. – С. 503-510.

## КВАНТОВАНИЕ КОЭФФИЦИЕНТОВ ЦИФРОВОГО КИХ-ФИЛЬТРА МЕТОДОМ ДИНАМИЧЕСКОГО ПРОГРАММИРОВАНИЯ

А.С. Бердышева,

ORCID: 0000-0002-8604-6346 e-mail: [iamanastasia\\_b@mail.ru](mailto:iamanastasia_b@mail.ru)

В.Н. Бугров

ORCID: 0000-0003-3220-9354 e-mail: [bug@rf.unn.ru](mailto:bug@rf.unn.ru)

Нижегородский государственный университет им. Н.И. Лобачевского

Приведено обоснование использования квантования коэффициентов цифрового КИХ-фильтра в концепции динамического программирования, как процесса пошагового квантования коэффициентов с их дискретной оптимизацией на каждом шаге по общей для всего процесса квантования целевой функции. Рассмотрен алгоритм пошагового динамического квантования методами целочисленного нелинейного программирования с учётом заданного масштабирования сигнала. Иллюстрируется эффективность применения данного подхода на примере динамического квантования коэффициентов каскадного КИХ-фильтра верхних частот с минимальной разрядностью представления целочисленных коэффициентов.

**Ключевые слова:** цифровой КИХ-фильтр, целочисленное нелинейное программирование, динамическое программирование, целевая функция.

## QUANTIZING THE COEFFICIENTS OF A FIR FILTER DYNAMIC PROGRAMMING METHOD

A.S. Berdysheva, V.N. Bugrov

National research Lobachevsky state university of Nizhny Novgorod

**Purpose:** To study the possibility and justification of using the quantization of the coefficients of the digital filter in the concept of dynamic programming, which allows to significantly reduce the functional error of quantization of the coefficients of the cascade FIR filter.

**Design / methodology / approach:** In classical analytical design, quantization errors in the coefficients of a digital FIR filter are considered unrecoverable, leading to distortion of the frequency characteristics and the need to scale its coefficients. In classical quantization, quantization errors determine the threshold for the bit representation of the coefficients when implementing a digital FIR filter on a particular digital platform. However, a significant decrease in the functional error can be ensured by using a "dynamic" coefficient quantization mechanism based on dynamic programming, as a section of optimal programming, in which the process of solving a general problem can be divided into separate stages (steps), and the results of the solution at one stage of the quantization process significantly affect for the next steps. R. Bellman's principle of optimality, which is the main content of dynamic programming, guarantees that the solution at any step is not locally better, but better from the point of view of the coefficient quantization process as a whole.

**Findings:** Dynamic quantization can significantly reduce the functional error of implementing the required characteristics of a low-bit FIR filter in comparison with classical statistical quantization. Modern minimization algorithms on a discrete grid with a given bit depth of parameter representation make it possible to solve such a dynamic problem reliably and efficiently, which makes it possible to significantly improve the quality of designed low-bit FIR filters

**Research limitations/implications:** To numerically solve the minimization problem in a quantized parametric space, an iterative algorithm of directed scanning on a deterministic grid with a discreteness equal to the number of binary digits that each variable is displayed in the code space is required, which fully complies with the requirements of the fixed-point format.

**Originality/value:** The novelty of the proposed quantization method is determined by the use of a dynamic coefficient quantization algorithm, which makes it possible to obtain a low-bit solution when designing high-speed FIR filters.

**Key words:** the FIR-filter, integer nonlinear programming, dynamic programming, criterion function

К основным эффектам конечной разрядности представления данных в системах цифровой фильтрации относятся погрешности характеристик, обусловленные квантованием коэффициентов цифрового фильтра. При представлении переменных в алгоритмах цифровой фильтрации в формате с фиксированной точкой (ФТ) все переменные цифрового алгоритма являются дробными вещественными числами, принадлежащими диапазону от  $-1$  до  $1-2^{-R}$ , где  $R$  – число значащих битов, используемых для двоичного представления числа [1, 2]. Таким образом, заданным числом двоичных разрядов можно реализовать только дискретный ряд определённых значений коэффициентов фильтра в интервале  $|x_i| < 1.0$ , где  $X(x_1, x_2 \dots x_i \dots x_n)$  – совокупный вектор коэффициентов цифрового КИХ-фильтра. Однако, аналитический расчёт коэффициентов, например, методом окна или частотной выборки, осуществляется в непрерывном  $n$ -мерном вещественном пространстве  $E^n$  с неограниченной точностью, поэтому на стадии практической реализации фильтра необходима процедура квантования его коэффициентов, которая в формате ФТ осуществляется следующим образом: для заданной длины битового представления  $R$  целочисленный код (квант) неквантованного коэффициента  $x_i \in E^n$  в варианте округления до ближайшего целого (round) равен целой части выражения в скобках:

$$x_i^{\text{int}} = \text{int}(x_i \cdot 2^R + 0.5), \quad (1)$$

Тогда вещественный квант данного коэффициента определяется так:

$$x_i^{\text{float}} = x_i^{\text{int}} / 2^R, \quad (2)$$

а ошибка квантования коэффициента составляет

$$\varepsilon_i = x_i - x_i^{\text{float}} \quad (3)$$

В рамках статистической модели [2] ошибки квантования коэффициентов считаются равномерно распределёнными случайными величинами с нулевым средним значением. В классическом аналитическом проектировании ошибки квантования коэффициентов цифрового фильтра считаются неустраняемыми, приводят к искажению частотных характеристик и необходимости масштабирования его коэффициентов.

Однако ошибку квантования коэффициентов можно устранить, осуществив дискретизацию параметрического пространства коэффициентов КИХ-фильтра только теми значениями, при которых ошибки квантования (3) равны нулю. При этом математическое задание квантованного параметрического пространства коэффициентов возможно как вещественным (2), так и целочисленным её кодом (1). КИХ-фильтры, использующие вещественную арифметику вычислений и дискретизацию коэффициентов, принято называть квантованными вещественными цифровыми фильтрами (ВЦФ), а фильтры, использующие целочисленную арифметику вычислений и дискретизацию параметров – целочисленными цифровыми фильтрами (ЦЦФ). Вариант целочисленного представления является более универсальным и практически значимым, так как ЦЦФ имеют минимальную вычислительную сложность и могут быть реализованы на любой цифровой платформе (MCU, DSP, FPGA) без наличия FPU–сопроцессора в структуре вычислителя [3]. При необходимости целочисленное решение легко преобразовать в вещественное квантованное решение формата ФТ, используя для этого их однозначную связь через соотношение (2). Формирование целочисленного кода часто называют нормализацией коэффициентов КИХ-фильтра, когда в дискретном  $n$ -мерном пространстве  $I^n_R$  формируется вектор целочисленных коэффициентов  $IX(ix_1, ix_2 \dots ix_i \dots ix_n)$ , квантованных с заданной разрядностью  $R$  их представления.

Ошибки квантования коэффициентов (3) вызывают, как уже сказано выше, соответствующую функциональную ошибку  $\Delta H_{(\omega)}^R$ , ошибку реализации требуемой характеристики КИХ-фильтра. Наиболее часто функциональная ошибка выполнения требований соответствует критерию среднеквадратичного отклонения (СКО) текущей характеристики от требуемой:

$$\Delta H_{(\omega)}^R = \frac{1}{p} \cdot \sum_{k=1}^p [Y_k^R(IX) - Y_k^T]^2 \quad (4)$$

где  $Y_k^R(\mathbf{IX})$  – текущее значение характеристики фильтра на  $k$ -ой дискретной частоте диапазона определения для заданной разрядности  $R$  представления коэффициентов, а  $Y_n^T$  – требуемое значение частотной характеристики КИХ-фильтра.

В классическом квантовании ошибки (3) и (4) определяют порог разрядного представления коэффициентов при реализации цифрового КИХ-фильтра на конкретной цифровой платформе. Приведённый алгоритм описывает классический «статический» механизм квантования коэффициентов фильтра, когда дальнейшее уменьшение разрядности вызывает уже недопустимую функциональную ошибку выполнения требований.

Однако существенное уменьшение функциональной ошибки можно обеспечить применением «динамического» механизма квантования коэффициентов на основе динамического программирования [4, 5], как раздела оптимального программирования, в котором процесс решения общей задачи может быть разбит на отдельные этапы (шаги), причём результаты решения на одном этапе процесса квантования существенно влияют на последующие шаги. Принцип оптимальности Р. Беллмана, являющийся основным содержанием динамического программирования, гарантирует, что решение на любом шаге не локально лучше, а лучше с точки зрения процесса квантования коэффициентов в целом.

Рассмотрим алгоритм динамического квантования на примере квантования коэффициентов нормализованного КИХ-фильтра, состоящего из каскадного соединения  $m$ -звеньев второго порядка ( $m=N/2$ , где  $N$  - порядок фильтра), с передаточной функцией [6]:

$$H(z) = \prod_{i=1}^m \frac{b_{0i} + b_{1i}z^{-1} + b_{2i}z^{-2}}{a_{0i}}, \quad (5)$$

Все коэффициенты передаточной функции (5) квантуются целочисленным кодом (1), а интервал их изменения определяется заданной разрядностью представления  $R$ . Разностное уравнение для одного звена такого целочисленного КИХ-фильтра имеет следующий вид:

$$y_n = (b_0x_n + b_1x_{n-1} + b_2x_{n-2})/a_0, \quad (6)$$

где  $x_n, y_n$  – входная и выходная целочисленные временные последовательности,  $a_0$  – нормирующий power-of-two коэффициент:

$$a_{0i} \in \{2^q\}, \quad q = \overline{0, R} \quad i = \overline{1, m} \quad (7)$$

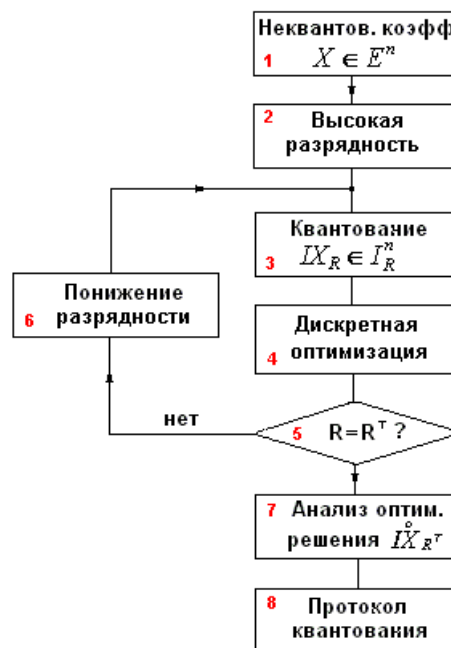


Рис. 1. Структура динамического квантования



На рис. 1 представлена общая блок-схема алгоритма динамического пошагового квантования целочисленных коэффициентов как последовательности поисковых задач с поэтапным понижением разрядности  $R$  представления коэффициентов фильтра вплоть до заданного минимального их представления  $R^T = R_{\min}$ .

На первом, стартовом этапе неквантованные вещественные коэффициенты  $X \in E^n$  сначала квантуются целочисленным кодом (1) со стартовой высокой разрядностью  $R_{\max} = 14$  бит (блоки 1–3), формируя тем самым дискретное целочисленное решение на квантованном множестве коэффициентов  $IX \in I_{R_{\max}}^n$  с соответствующей функциональной ошибкой квантования (4). Затем на этом дискретном множестве ставится задача оптимизации дискретных коэффициентов данной разрядности по общей для всех этапов динамического квантования целевой функции (4), как задачи целочисленного нелинейного программирования (блок 4), которая для каскадного КИХ-фильтра с передаточной функцией (5) записывается так:

$$F^o(IX_R) = \min F(IX_R) \quad IX_R \in I_R^{3m} \quad (8)$$

$$-2^R < b_{di} < 2^R \quad d = \overline{0, 2} \quad i = \overline{1, m}, \quad (9)$$

$$a_{0i} \in \{2^q\}, \quad q = \overline{0, R} \quad i = \overline{1, m}, \quad (10)$$

$$K_i^{\min} \leq |K_i(e^{j\omega T})| \leq K_i^{\max} \quad i = \overline{1, m}, \quad (11)$$

где  $F(IX_R)$  - целевая функция,  $d$  - индекс коэффициента передаточной функции (5),  $K_i^{\min}$ ,  $K_i^{\max}$  - допустимые границы изменения коэффициента усиления  $i$ -го звена.

Для численного решения экстремальной задачи (8) в квантованном параметрическом пространстве используется итеративный алгоритм направленного сканирования на детерминированной сетке с дискретностью, равной числу двоичных разрядов  $R$ , которыми в кодовом пространстве отображается каждая  $i$ -ая переменная [6, 7], что полностью соответствует требованиям формата ФТ. Характерными особенностями поискового алгоритма является высокая надёжность отделения глобального экстремума, малые потери на поиск, эффективная работа в пространстве высокой размерности, а также отсутствие априори настраиваемых параметров. Далее в блоке 5 производится проверка соответствия текущей разрядности квантования  $R$  с требуемой минимальной разрядностью  $R^T$  и в случае несоответствия осуществляется понижение разрядности представления коэффициентов (блок 6) с последующим их квантованием и оптимизацией на новом дискретном множестве (блоки 3 и 4) также по общей целевой функции ошибки (4). Таким образом, решение на текущем этапе динамического квантования является стартовой точкой этапа последующего. Цикл повторяется до достижения требуемой минимальной разрядности представления коэффициентов, после чего осуществляется выявление оптимального малоразрядного решения, обеспечивающего допустимый уровень ухудшения функциональных характеристик фильтра, и проводится анализ этого решения во временной и частотной областях.

Рассмотрим пример динамического квантования коэффициентов исходного синтезированного КИХ-фильтра верхних частот при целочисленном описании его квантованного варианта. В итоге необходимо определить минимальную разрядность квантования коэффициентов, при которой выполняются следующие требования к функциональным показателям КИХ-фильтра:

1. Полоса пропускания: 450 – 1000 Гц
2. Коэффициент передачи в полосе пропускания: 0 дБ с допуском  $\pm 1$  дБ
3. Уровень подавления на частотах ниже 200 Гц: 40 дБ
4. Нелинейность фазы в полосе:  $10^\circ$
5. Порядок КИХ-фильтра: 56
6. Масштабирование коэффициентов передачи звеньев в интервал  $\{0.5 - 2\}$
7. Частота дискретизации: 2 кГц

Необходимо отметить, что синтез исходного неквантованного фильтра верхних частот (ФВЧ) осуществлялся как по критерию требуемой частотной селективности, так и по критерию обеспечения возможной линейности фазы в полосе пропускания фильтра. Такие же требования

сохранялись и при динамическом его квантовании, когда целевой функцией являлась взвешенная СКО выполнения функциональных требований к АЧХ и ФЧХ квантуемого КИХ-фильтра

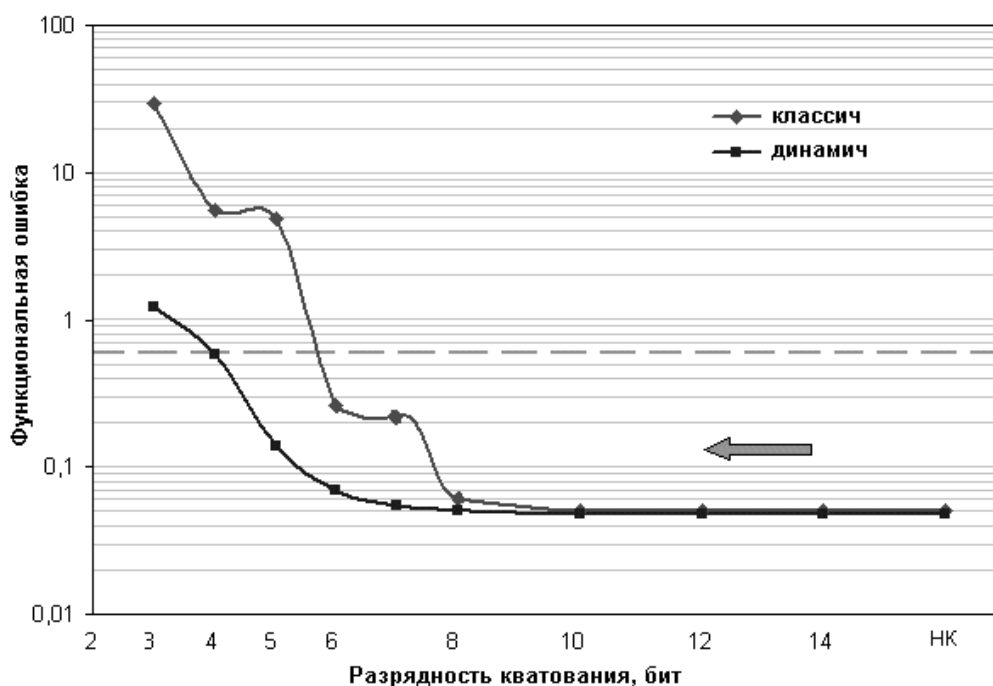
$$F(\mathbf{IX}_R) = \beta_1 f_{\text{АЧХ}}(\mathbf{IX}_R) + \beta_2 f_{\text{ФЧХ}}(\mathbf{IX}_R), \quad (12)$$

где частные целевые функции  $f_{\text{АЧХ}}(\mathbf{IX}_R)$  и  $f_{\text{ФЧХ}}(\mathbf{IX}_R)$ , обеспечивающие выполнение требований амплитудной селекции фильтра и фазовой его линейности, определялись соответствующей функциональной ошибкой (4). График требуемой АЧХ фильтра приведен на рис. 3а (выделен пунктиром).

Осуществим сначала классическое квантование коэффициентов данного фильтра в варианте округления по соотношению (1) последовательно до минимальной разрядности их представления 3 бита, фиксируя функциональную ошибку (13) реализации требуемой АЧХ и ФЧХ фильтра в таблице 1. Из таблицы видно, что разброс квантования по функциональной ошибке превосходит два порядка, поэтому график соответствующей зависимости для наглядности представлен в логарифмическом масштабе (рис. 2). Необходимо заметить, что далее в таблице и графиках разрядность квантования целочисленных коэффициентов указана уже с учётом знакового бита.

Ошибки статистического и динамического квантования Таблица 1.

СКО	Разрядность квантования, бит									
	НК	14	12	10	8	7	6	5	4	3
$\Delta\mathbf{H}$ клас	0.050	0.051	0.051	0.053	0.06	0.21	0.26	4.82	5.50	29.6
$\Delta\mathbf{H}$ дин	0.050	0.050	0.050	0.051	0.05	0.054	0.07	0.14	0.58	1.51



**Рис. 2. Ошибки классического и динамического квантования**

Осуществим теперь динамическое квантование коэффициентов исходного фильтра по приведённому выше алгоритму и целевой функции (13) также последовательно от максимальной до минимальной 3-битовой разрядности их представления, фиксируя функциональную ошибку  $\Delta\mathbf{H}$  дин реализации требуемой АЧХ и ФЧХ в таблице 1. График динамической ошибки квантования также представлен на рис. 2.

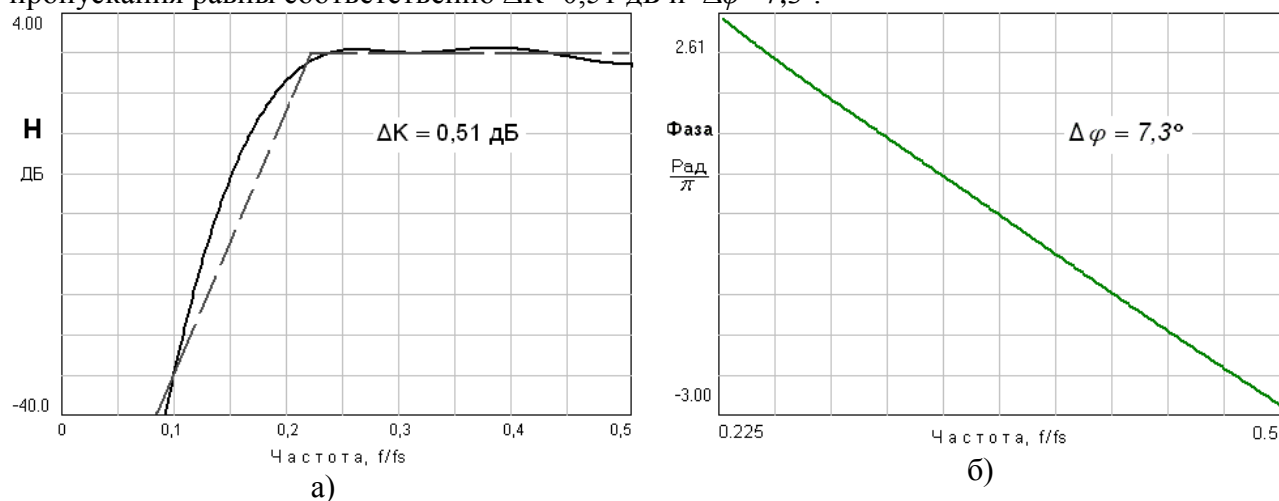
Из таблицы 1 и приведённых графиков видно, что функциональная ошибка классического квантования резко возрастает начиная с уровня 10–битового представления коэффициентов КИХ-фильтра, в то время как ошибка динамического квантования имеет тенденцию к росту только

начиная с уровня 6-битового их представления. Также можно отметить, что допустимому уровню 0,6 (отображен из рис. 2 пунктиром) выполнения спецификационных требований по функциональной ошибке квантования соответствует только 6-битовое решение при классическом квантовании, в то время как при динамическом квантовании для представления коэффициентов вполне достаточно 4-х бит.

Оптимальные коэффициенты 4-битового ФВЧ Таблица 2.

Звено	$b_0$	$b_1$	$b_2$	$K_{\max}$
1	0	0	-7	0,9
2	7	-7	-2	1,5
3	-1	-2	-7	1,3
4	3	7	-5	1,3
5	7	-3	-1	1,1
6	-2	-7	7	1,5
7	7	-7	-2	1,5
8	7	-7	-2	1,5
9	6	4	1	1,4
10	7	-1	0	1,0
11	7	-7	-2	1,5
12	7	-7	-2	1,5
13	7	-1	0	1,0
14	7	-7	-2	1,5
15	7	-1	0	1,0
16	7	1	1	1,1
17	7	0	0	0,9
18	7	0	0	0,9
19	7	2	1	1,3
20	7	0	0	0,9
21	7	0	0	0,9
22	7	-7	-6	1,9
23	0	-7	2	1,1
24	5	2	2	1,1
25	7	0	0	0,9
26	7	0	0	0,9
27	7	0	0	0,9
28	7	0	0	0,9

В таблице 2 приведены оптимальные значения этих целочисленных 4-битовых коэффициентов передаточной функции по динамическому квантованию ФВЧ со значимостью  $\beta_1=5\beta_2$  требований к АЧХ и ФЧХ, а также достигнутый уровень его каскадного масштабирования. Графики частотных характеристик представлены на рис. 3. По критерию СКО неравномерность АЧХ и ФЧХ в полосе пропускания равны соответственно  $\Delta K=0,51$  дБ и  $\Delta\varphi=7,3^\circ$ .



### Рис. 3. Характеристики 4-битового ФВЧ а) АЧХ, б) ФЧХ

Таким образом, все функциональные требования при динамическом квантовании коэффициентов 4-битового нерекурсивного ФВЧ 56-го порядка были выполнены. Экспериментальные измерения полностью подтвердили частотные характеристики отквантованного КИХ-фильтра.

#### Выводы

Динамическое квантование позволяет существенно уменьшить функциональную ошибку реализации требуемых характеристик малоразрядного КИХ-фильтра в сравнении с классическим статистическим квантованием. Современные алгоритмы минимизации на дискретной сетке с заданной разрядностью представления параметров позволяют решать такую динамическую задачу надёжно и эффективно, что даёт возможность существенно повысить качество проектируемых малоразрядных КИХ-фильтров.

#### Библиографический список

1. **Айфичер, Э., Джервис Б.** Цифровая обработка сигналов: практический подход. М.: Издательский дом "Вильямс", 2004.
2. **Рабинер, Л., Гоулд, Б.** Теория и применение цифровой обработки сигналов. - М.: Мир, 1978.
3. **Бугров, В.Н.** Дискретный синтез минимально-фазовых и линейно-фазовых цифровых БИХ-фильтров. // Компоненты и технологии, № 10, 2019.
4. **Вентцель, Е.С.** Исследование операций. Задачи, принципы, методологии. – М.: Наука, 1988, -208 с.
5. **Есипов, Б.А.** Методы оптимизации и исследования операций. – Самара: Изд-во Самар. гос. аэрокосм. ун-та, 2007, 180 с.
6. **Бугров, В.Н.** Целочисленные цифровые фильтры. Монография. Saarbrücken: Lambert Academic Publishing, 2017.
7. **Воинов, Б.С., Бугров, В.Н., Воинов, Б.Б.** Информационные технологии и системы: поиск оптимальных, оригинальных и рациональных решений. М.: Наука, 2007.

**ДИСКРЕТНЫЙ СИНТЕЗ АКТИВНЫХ ФИЛЬТРОВ ЗАЩИТЫ  
ОТ НАЛОЖЕНИЯ СПЕКТРОВ****Ю.М. Макарова**ORCID: 0000-0002-6107-1211 e-mail: [julietta.mak@mail.ru](mailto:julietta.mak@mail.ru)**В.Н. Бугров**ORCID: 0000-0003-3220-9354 e-mail: [bug@rf.unn.ru](mailto:bug@rf.unn.ru)

Нижегородский государственный университет им. Н.И. Лобачевского

В докладе рассматривается проектирование активного полиномиального полосового фильтра поисковыми методами нелинейного математического программирования в дискретном пространстве параметров. Проектирование осуществляется с учётом реальных моделей операционных усилителей путём их макро моделирования в частотной области. Рассмотрен пример решения задачи синтеза активного фильтра защиты от наложения спектров гидроакустического приёмного тракта. Проведён анализ его теоретических и экспериментальных характеристик.

**Ключевые слова:** активный фильтр, наложение спектров, дискретное программирование, целевая функция.

**DISCRETE SYNTHESIS OF ACTIVE FILTERS  
PROTECTING FROM SPECTRUM OVERLAP****J.M. Makarova, V.N. Bugrov**

National research Lobachevsky state university of Nizhny Novgorod

**Purpose:** The authors design an active polynomial filter using discrete mathematical optimization. The filter provides protection against spectrum overlap for hydroacoustic receiving station.

**Design / methodology / approach:** The search methods of non-linear mathematical programming in a discrete parameter space (resistor and capacitor values) is an alternative approach to active filter design. The idea of this methodology is to link the design task to a clear invariant mathematical feature – the extremum of the filter aggregate quality function, called objective function. This function is formed on the basis of the required filter characteristics. The synthesis problem is reduced to minimizing objective function, i.e. finding the coordinates of the global extremum. Minimization is made by search methods. In this paper the readers are introduced to a solution of the synthesis problem obtained with the BARC search program, developed at the radio engineering department by the authors. This package is intended for a multifunctional search design of active devices with an arbitrary structure. At the design stage op amps are replaced by the macromodels in a frequency domain, whose parameters are also computed with search methods using the BARC package. Verification of the found solution was carried out by modeling and measuring the characteristics of a real sample. It is confirmed that the requirements are fulfilled.

**Findings:** In this paper the design of an active filter using discrete optimization is presented. Set of the requirements are satisfied, namely efficient signal suppression above the Nyquist frequency and low-frequency noise suppression, minimal amplitude distortion in the pass-band, steep cutoff and minimal width of the transition bands, wide dynamic range, discretization of the filter component values, minimal dimensions and weight. Highly reliable solution is obtained despite the complex multiextremal form of the objective function. The electronic circuit is implemented in production.

**Research limitations/implications:** The requirement of phase linearity was not set within this design problem. When computing the filter complex transfer function by the node voltage method, instead the real op amp, its macromodel is considered, which, however, reflects the characteristics of the op amp with sufficient approximation accuracy for this problem.

**Originality/value:** Search methods of discrete optimization in an active filter design are a promising and advanced alternative to their traditional analytical calculations. The main difference is that filter synthesis is carried out by a direct search of the component values in a multidimensional discrete space, that makes it possible to design an active filter with any topology, to control easily the priority of the functional characteristics during the filter synthesis, to scale signal amplification in the filter sections and others.

**Key words:** active filter, spectrum overlap, discrete programming, objective function

## Введение

Цифровая обработка гидроакустических сигналов включает в себя преобразование аналогового сигнала в цифровой с помощью аналого-цифрового преобразователя (АЦП). При этом для предотвращения появления ложных сигналов из-за эффекта наложения спектров дискретизированных сигналов [1], необходимо ограничивать спектр сигнала сверху перед оцифровкой аналоговым фильтром. Частота среза такого активного фильтра защиты от наложения спектров (ФЗНС) устанавливается равной частоте Найквиста, т.е. половине частоты дискретизации ( $f_s/2$ ). Наиболее часто в системе подводной связи используются фазоманипулированные сигналы. Однако анализ спектра на входе гидроакустического приёмного тракта показывает, что характерной его особенностью является также наличие шумовых сигналов от близлежащих гидроакустических объектов в низкочастотной области (до 50 – 100 Гц), мощность которых может на порядки превышать уровень полезного гидроакустического сигнала (рис. 1).

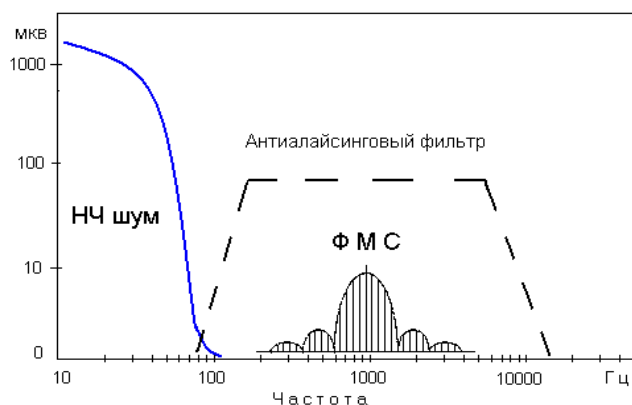


Рис. 1. Сигналы на входе приёмной станции

Наличие такой мощной низкочастотной помехи существенно снижает динамический диапазон гидроакустического приёмного тракта и может вызвать серьезные нарушения работы АЦП при оцифровке аналогового сигнала. Поэтому весьма желательно построение ФЗНС с функцией высокого подавления подобной низкочастотной помехи, то есть осуществлять проектирование ФЗНС как полосового фильтра, обеспечивающего требования как подавления всех сигналов выше частоты Найквиста, так и подавления шумовой низкочастотной помехи. Таким образом, основными требованиями к проектированию фильтра защиты являются:

- высокий уровень подавления сигналов выше частоты Найквиста;
- высокий уровень подавления низкочастотной шумовой помехи;
- обеспечение минимальных искажений в полосе пропускания фильтра;
- высокий коэффициент прямоугольности АЧХ, минимальная ширина переходных зон;
- широкий динамический диапазон ФЗНС;
- дискретизация параметров фильтра (номиналов резисторов и конденсаторов);
- минимальные габариты и вес.

Для обработки гидроакустических сигналов в области низких и ультранизких частот выполнить данные требования возможно только активным RC фильтром (ARC-фильтром) достаточно высокого порядка.

В настоящее время среди технологий проектирования ARC-фильтров преобладает аналитический расчёт [2-6], который имеет следующие недостатки: принципиальная невозможность многофункционального синтеза, ограниченный выбор топологии построения проектируемого ARC-фильтра, отсутствие возможности получения проектного решения в дискретном пространстве, возможность использовать лишь идеальные модели операционных усилителей и др.

Альтернативным направлением, не имеющим вышеперечисленных недостатков, является поисковое проектирование активных фильтров в дискретном пространстве параметров (значений резисторов и емкостей). Синтез технического решения при этом сводится к задаче нелинейного

математического программирования [7, 8]. Минимизация целевой функции при этом обычно осуществляется поисковыми методами [9].

### Постановка задачи

В предлагаемой статье авторы знакомят читателей с поисковым решением задачи синтеза с помощью эксклюзивной поисковой программы BARC, предназначенной для многофункционального поискового проектирования ARC-устройств произвольной структуры в непрерывном или дискретном пространстве параметров.

Рассмотрим работу данной программы на примере решения конкретной задачи дискретного синтеза полосового ФЗНС гидроакустического приёмного тракта по следующим функциональным требованиям:

1. Полоса пропускания фильтра 180 – 800 Гц
2. Коэффициент передачи  $K_u=1,0$  (0 дБ)
3. Неравномерность  $K_u$  в полосе  $\pm 0,5$  дБ
4. Коэффициент подавления на частотах 50 и 3000 Гц не хуже  $-40$  дБ
5. Фазовые искажения в полосе пропускания не более  $60^\circ$
6. Дискретизация параметров фильтра по ряду E12 (10%)

График требуемой АЧХ антиалайсингового фильтра представлен на рис. 2.

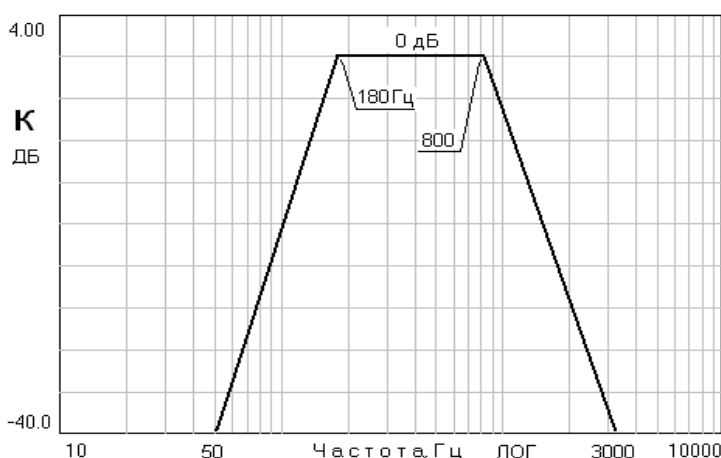


Рис. 2. Требуемая амплитудно-частотная характеристика ААФ

### Макромоделирование операционного усилителя

На первом этапе проектирования необходимо выбрать операционный усилитель (ОУ), на базе которого будет реализована схема ФЗНС, и сформировать его линейную макромодель замещения. Для реализации приведённого выше технического задания вполне подходит операционный усилитель MAX44241 от фирмы Maxim, частотный диапазон которого с большим запасом удовлетворяет требованиям, а зависимость коэффициента передачи от частоты довольно простая и имеет только одну ярко выраженную точку перегиба. Такая зависимость с высокой точностью может быть смоделирована постановкой только одного полюса передаточной функции макромодели (рис. 3). Однополюсная модель требует минимальных ресурсов оперативной памяти ЭВМ при моделировании ОУ, так как имеет простую топологию и всего один внутренний узел.

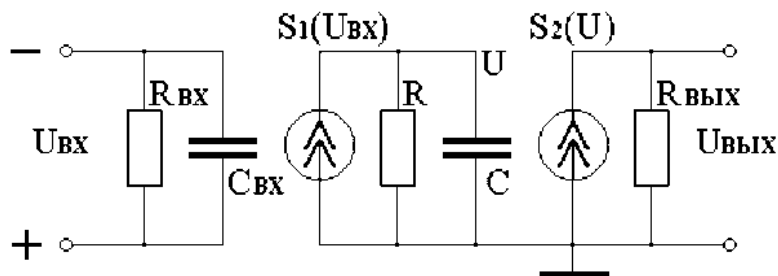


Рис. 3. Однополюсная макромодель ОУ MAX44241

Таким образом, данная линейная макромоделль замещения имеет такой же входной и выходной импедансы и такую же зависимость передаточной функции от частоты, как и реальный ОУ. Точка перегиба на частоте 1 Гц соответствует отрицательному действительному полюсу и смоделирована источником тока  $S_1$ , управляемым напряжением  $U_{вх}$ , с крутизной преобразования  $1 \text{ а/в}$ , резистором  $R=300 \text{ кОм}$  и емкостью  $C=1 \text{ мкФ}$ . Остальные параметры макромоделли имеют следующие значения:  $R_{вх}=1 \text{ ГОм}$ ,  $C_{вх}=2 \text{ пФ}$ ,  $R_{вых}=300 \text{ Ом}$ , крутизна преобразования  $1,6 \text{ а/в}$  для источника тока  $S_2$ , управляемого напряжением  $U$ . Часть параметров данной макромоделли были взяты из паспортных данных ОУ ( $R_{вх}$ ,  $C_{вх}$ ,  $R_{вых}$ ), а остальные параметры оптимизированы тем же пакетом VARC по требуемой паспортной АЧХ ОУ MAX44241.

### Топология построения фильтра

На втором этапе необходимо выбрать топологию построения проектируемого активного фильтра. Из опыта практической разработки известно, что область применения простой каскадной реализации фильтра на звеньях одинаковой структуры имеет ограниченные возможности выполнения сложных функциональных требований. Многие недостатки такого построения ФЗНС могут быть устранены применением связно-каскадной реализации, состоящей из каскадного соединения звеньев первого или второго порядков разной структуры, охваченных попарно цепью отрицательной обратной связи [3, 8]. Чувствительность к изменениям параметров в таких фильтрах намного меньше, а возможность реализации требуемых характеристик существенно возрастает. Кроме того, за счет каскадирования связанных блоков в таких фильтрах удается уменьшить паразитное прохождение сигналов высокой частоты на выход фильтра, имеющее место в большинстве ARC-фильтров с многопетлевой обратной связью.

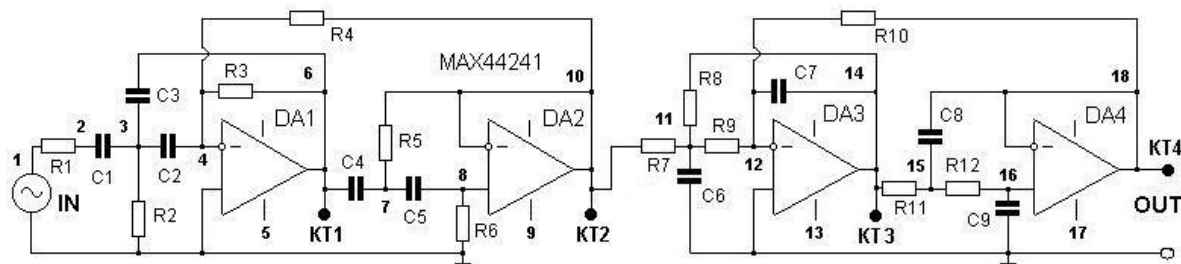


Рис. 4. Топология полиномиального активного полосового ФЗНС

Остановим свой выбор на структуре полиномиального ФЗНС восьмого порядка на связно-каскадном соединении двух звеньев ФВЧ и двух звеньев ФНЧ с использованием межкаскадной отрицательной обратной связи. Применение звеньев различной частотной селекции дает возможность независимого управления крутизной низкочастотного и высокочастотного срезов АЧХ простым добавлением дополнительных звеньев в линейку ФНЧ или ФВЧ проектируемого фильтра.

Передаточная функция полиномиальных фильтров определяется выражением

$$K(p) = \frac{K_0}{v(p)}, \quad (1)$$

где  $v(p)$  – полином Гурвица порядка  $n$ , все корни которого лежат в левой полуплоскости комплексной частоты  $p = \sigma + j\omega$ , а постоянный множитель  $K_0$  определяет коэффициент передачи фильтра на нулевой частоте. Таким образом, полиномиальный фильтр не имеет нулей коэффициента передачи, а все полюса его конечны. Это обеспечивает более высокую фазовую линейность и устойчивость полиномиальной линейной динамической системы, а также более простую её реализацию по сравнению с линейными цепями с конечными нулями коэффициента передачи. На рис. 4 приведена топология такого фильтра для переменной составляющей тока. При анализе в частотной области выбранная схема построения фильтра трактуется как линейная стационарная цепь с замещением реальных операционных усилителей их линейными макромоделями и численным расчётом комплексного коэффициента передачи в заданном частотном диапазоне методом узловых потенциалов (цифрами на схеме обозначены номера независимых узлов). Это позволяет моделировать и синтезировать данным пакетом активные фильтры самой разнообразной топологии построения.



### Дискретный синтез ARC-фильтра

При синтезе активного полосового фильтра редактор функциональных характеристик программы BARC осуществляет ввод в графическом режиме требуемой АЧХ фильтра и формирует целевой функционал экстремальной задачи синтеза по критерию минимума среднеквадратичного отклонения

$$F(X) = \sqrt{\frac{1}{P} \cdot \sum_{n=1}^P [Y_n(X) - Y_n^T]^2}, \quad (2)$$

где  $Y_n(X)$  – текущее значение АЧХ фильтра на  $n$ -ой дискретной частоте диапазона определения, а  $Y_n^T$  – требуемое значение частотной характеристики.

Относительно целевой функции (2) задача синтеза полиномиального полосового активного фильтра выглядит так:

$$F^o(X^o) = \min F(X) \quad X \in DX \quad (3)$$

$$DX : \begin{cases} R_i \in E12 \\ C_i \in E12 \\ 100 \text{ Ом} \leq R_i \leq 1 \text{ МОм} \\ 100 \text{ нФ} \leq C_i \leq 4,42 \text{ нФ} \end{cases}, \quad (4)$$

$$0,6 \leq |K_i(j\omega)| \leq 1,2 \quad i = \overline{1,3} \quad (5)$$

Экстремальная задача дискретного программирования (3) записана относительно многомерного пространства проектирования  $DX$ , дискретизация которого осуществлялась по ряду  $E12$ . Ограничения (4) при этом определяли границы изменения дискретных варьируемых параметров. В данной задаче под оптимизацию были поставлены все 20 внешних параметров ФЗНС, кроме параметров макромоделей ОУ и сопротивления генератора входного гидроакустического сигнала  $R_T=R_1=600 \text{ Ом}$ . Функциональные ограничения (5) масштабировали коэффициенты усиления каскадов (по контрольным точкам КТ1 – КТ3) в заданный интервал, обеспечивая устойчивость работы проектируемого фильтра в широком динамическом диапазоне входных сигналов.

Для численного решения экстремальной задачи (3) в программе используется алгоритмический комплекс многомерной дискретной минимизации, в состав которого входит эффективный алгоритм глобальной минимизации на сетке кода Грея [9], адаптированный к поиску решений в режиме дискретного представления многомерной области проектирования стандартными рядами  $E6-E192$ .

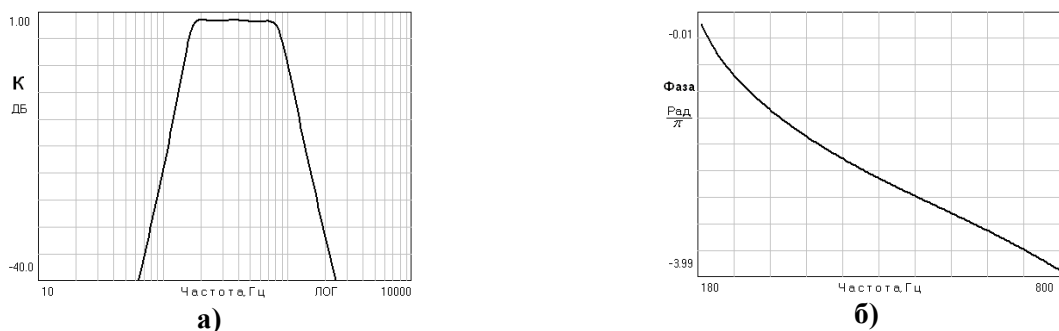
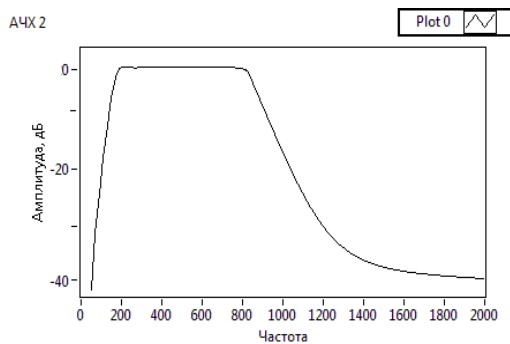
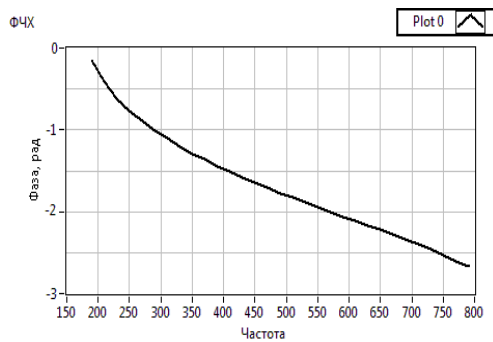


Рис. 5. Характеристики ФЗНС по синтезу: а) АЧХ б) ФЧХ



а)



б)

Рис. 6. Измеренные характеристики ФЗНС: а) АЧХ б) ФЧХ

Время решения данной задачи на стандартном персональном компьютере не превышало 10 минут. В таблице 1 приведены значения оптимальных дискретных параметров, а в таблице 2 – функциональных показателей синтезированного ФЗНС по результатам компьютерного синтеза и экспериментальных измерений. На рис. 5, 6 представлены графики его характеристик. Измерение (рис. 6) осуществлялось на реальном сигнале с помощью автоматизированной панорамной измерительной системы, разработанной в среде виртуальных приборов LabVIEW.

Дискретные параметры фильтра

Таблица 1

Резисторы				Конденсаторы			
R1	600 Ом	R7	12 кОм	C1	3,9 нФ	C7	2,7 нФ
R2	39 кОм	R8	82 кОм	C2	15 нФ	C8	2,2 нФ
R3	100 кОм	R9	27 кОм	C3	5,6 нФ	C9	100 пФ
R4	68 кОм	R10	820 кОм	C4	2,2 нФ		
R5	47 кОм	R11	180 кОм	C5	3,3 нФ		
R6	820 кОм	R12	820 кОм	C6	47 нФ		

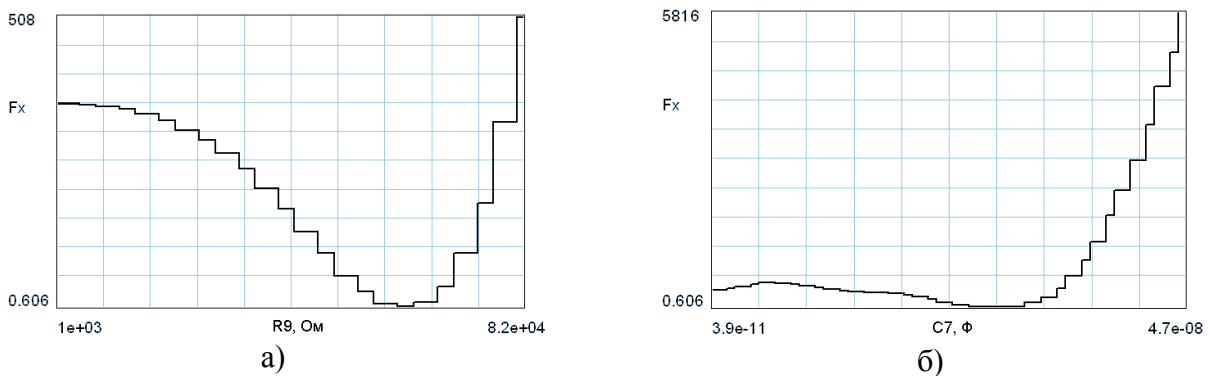
Показатели антиалайсингового фильтра

Таблица 2

Функциональные показатели фильтра	Синтез на ЭВМ	Измерение
1. Полоса пропускания, Гц	180 – 800	178 – 810
2. Затухание на частотах 50 и 3000 Гц не менее, дБ	-40	-42
3. Усиление в полосе, дБ	0	0,072
4. Неравномерность усиления, дБ	0,42	0,65
5. Фазовые искажения в полосе, град	54	47

Как видно из таблиц и приведённых графиков все требования по функциональным показателям и характеристикам полосового полиномиального ФЗНС были выполнены с высокой точностью. При этом в ходе синтеза был достигнут высокий уровень масштабирования сигнала. Максимальные значения коэффициентов усиления каскадов (по контрольным точкам КТ1, КТ2 и КТ3) находятся в границах заданного интервала  $\{0,6 - 1,2\}$ .

Исследуем профиль целевого функционала в точке оптимума путём построения его координатных разрезов. Графики разреза по параметрам  $R_9$  и  $C_7$  второго каскада (рис. 7) показывают, что целевые функции в задачах синтеза ARC-фильтров имеют сложный, полимодальный характер. Дискретная минимизация таких функций является весьма непростой задачей. Тем не менее, разработанный программно-алгоритмический поисковый комплекс пакета BARC 2 успешно справился с этой задачей, показав высокую надёжность и эффективность.



**Рис. 7. Координатный разрез целевой функции по параметру R9 (а) и параметру C7 (б)**

Особенностью найденного решения является значительная нелинейность ФЧХ. Эта нелинейность может быть существенно уменьшена использованием ФЗНС с гауссовой огибающей АЧХ, синтез которых является предметом дальнейших исследований.

### Заключение

Требования к селективным свойствам активных фильтров защиты от наложения спектров в гидроакустических приёмных трактах непрерывно возрастают, что неизбежно вызывает необходимость совершенствования методологий их моделирования и синтеза. Поисковые методы дискретного программирования в приложении к задачам проектирования АРС-фильтров являются перспективной и современной альтернативой традиционному аналитическому их расчёту. Принципиальное отличие заключается в том, что в данном случае для удовлетворения требуемого функционирования ААФ на стадии его проектирования осуществляется прямой поиск параметров фильтра в многомерном дискретном пространстве проектирования. Критерием поиска эффективного решения является соответствие совокупного текущего функционирования активного фильтра его требуемому функционированию. Это даёт возможность существенно повысить качество проектируемых ААФ, сократить время их разработки.

Из материалов, приведённых в данной статье видно, что поисковый синтез активных фильтров методом дискретного нелинейного программирования, позволяет:

- 1) синтезировать активные фильтры любой топологии с численным расчётом комплексного коэффициента передачи и его частотных характеристик методом узловых потенциалов;
- 2) осуществлять синтез фильтра по совокупности требуемых характеристик, причём можно легко управлять приоритетом функциональных характеристик в процессе синтеза фильтра;
- 3) осуществлять поисковый синтез проектного решения с учётом особенностей реальных операционных усилителей адекватным макро моделированием в частотной области;
- 4) достичь высокой надёжности отыскания эффективного решения экстремальной задачи синтеза ААФ благодаря глобальной модельной идеологии поискового алгоритма;
- 5) обеспечить необходимое масштабирование сигнала в каскадных структурах фильтров непосредственно в процессе поискового их синтеза. Здесь нет необходимости применения косвенных приёмов масштабирования усиления;
- 6) дискретизировать пространство параметров стандартными рядами E6 – E192, по которым радиоэлектронные компоненты выпускаются промышленностью, что позволяет получать проектные решения, реализация которых не вызывает никаких практических затруднений.

### Библиографический список

1. **Айфичер, Э.** Цифровая обработка сигналов: практический подход. / Э. Айфичер, Б. Джервис // М.: Издательский дом "Вильямс". 2004. 992 с.
2. **Рентюк, В.** Проектирование активных фильтров в Analog Filter Wizard 2.0. // Компоненты и технологии. Москва. 2013. № 6. С. 130-136.

3. **Джонсон, Д.** Справочник по активным фильтрам. / Д. Джонсон, Дж. Джонсон, Г. Мур // М.: Энергоатомиздат. 1983. 128 с.
4. **Moschytz, G.S.** Active Filter Design Techniques. // Analog Circuit Theory and Filter Design in the Digital World. Springer. 2019
5. **Ульянов, А. В.** Проектирование активных фильтров с использованием FilterLab 2.0 / А. В. Ульянов, М. В. Линник // Технические науки: проблемы и перспективы: материалы V Междунар. науч. конф. (г. Санкт-Петербург, июль 2017 г.). – Санкт-Петербург: Свое издательство, 2017. С. 32-37. URL: <https://moluch.ru/conf/tech/archive/231/12550>
6. **Буренко, Е.А.** Проектирование активных частотно-избирательных фильтров. / С.В. Нестеров, Е.А. Буренко // Труды МАИ. Выпуск № 101. 2018
7. **Мину, М.** Математическое программирование. Теория и алгоритмы. / М.: Наука. 1990. 488 с.
8. **Бугров, В.Н.** Поисковое проектирование активных антиалайсинговых фильтров / В.Н. Бугров, Ю.М. Макарова // Проектирование и технология электронных средств. Владимир. 2017. № 1. С. 35-41
9. **Воинов, Б.С.** Информационные технологии и системы: поиск оптимальных, оригинальных и рациональных решений. / Б.С. Воинов, В.Н. Бугров, Б.Б. Воинов // М.: Наука. 2007. 730 с.

**УСТРОЙСТВО МЕХАНИЧЕСКОГО СКАНИРОВАНИЯ  
БОРТОВОГО РАДИОЛОКАТОРА С МАГНИТОМЕТРИЧЕСКОЙ СИСТЕМОЙ  
ВСТРОЕННОГО КОНТРОЛЯ УГЛОВОГО ПОЛОЖЕНИЯ ЛУЧА**

**А.В. Голубцов**

ORCID: 0000-0002-0736-8532 e-mail: [myLetterBox@yandex.ru](mailto:myLetterBox@yandex.ru)

**Ю.М. Куликов**

ORCID: 0000-0002-8549-8906 e-mail: [umkulikov@yandex.ru](mailto:umkulikov@yandex.ru)

**Н.А. Макарычев**

ORCID: 0000-0002-7981-0756 e-mail: [nickmak52@gmail.com](mailto:nickmak52@gmail.com)

**С.В. Марьевский**

ORCID: 0000-0002-8763-1336 e-mail: [777serzhant888@mail.ru](mailto:777serzhant888@mail.ru)

**Е.А. Михалицын**

ORCID: 0000-0003-3144-3704 e-mail: [mihalitsynea@gmail.com](mailto:mihalitsynea@gmail.com)

**А.В. Назаров**

ORCID: 0000-0002-0261-1724 e-mail: [nazarov52@mail.ru](mailto:nazarov52@mail.ru)

Филиал РФЯЦ-ВНИИЭФ «НИИИС им.Ю.Е. Седакова»

В докладе предложена новая конструкция устройства углового позиционирования бортового радиолокатора миллиметрового диапазона длин волн с магнитометрической системой встроенного контроля положения луча. Разработаны математическая модель, описывающая зависимость углов сканирования от угловых параметров серводвигателей, и алгоритм управления системой сканирования. В конструкции используются сервоприводы с автоматическим контролем положения, что позволяет в первом приближении не учитывать в математической модели триботехнические потери. Проведен анализ различных законов сканирования радиолокатора. Выбран оптимальный для разработанной конструкции закон сканирования. Предложены способ секторального азимутального отслеживания текущего положения луча с фиксацией нулевого отсчета и косвенный метод контроля текущего угла элевации.

**Ключевые слова:** устройство углового позиционирования радиолокатора, датчик магнитного поля, закон сканирования.

**THE AIRBORNE RADAR MECHANICAL SCANNING UNIT WITH BEAM  
ANGLE POSITION BUILT-IN TEST SYSTEM BASED ON MAGNETIC FIELD SENSORS  
A.V. Golubzov, U.M. Kulikov, N.A. Makarychev, S.V. Marevskiy, E.A. Mihalitsyn, A.V. Nazarov**

Branch of federal state unitary enterprise «Russian federal nuclear center - All-russian scientific-research institute of experimental physics» «Scientific-research institute of measuring systems  
n.a. Yu.E. Sedakov»

**Purpose:** The authors provide the new design construction of millimeter wave airborne radar beam angle positioning unit which includes built-in test system based on magnetic field sensors to control beam movements.

**Design / methodology / approach:** The mathematical model for controlling mechanical system is presented here. The paper contains new information to concern of the sectorial azimuthally control method which could be used for zero-position fixing and the continuous indirect approach of elevation beam scanning angle measuring.

**Findings:** The paper finds out optimal scanning law for designed unit construction. It also makes conclusions about advantages and disadvantages of each scanning laws based on chosen quality metrics.

**Research limitations/implications:** Research model is limited to no friction case. Nevertheless, the obtained results could be further used in real radar subsystems with automatic control of servo motors positioning.

**Originality/value:** The value of the whole paper is the new design construction of angle positioning and scanning unit which allows to setup desirable angles position and perform elevation control and azimuthally tracking.

**Key words:** angle positioning unit of radar, magnetic field sensor, scanning law.

## Введение

Несмотря на бурное развитие бортовых радиолокационных систем с электронным сканированием, механические системы находят применение не только в качестве дополнительного способа расширения диапазона углового обзора [1], но и автономно в радиолокаторах миллиметрового диапазона при смягченных требованиях к скоростям сканирования и дальности обнаружения. Кроме того, механические системы позиционирования используются для стабилизации бортовых КВЧ метеорадаров [2]. Поэтому разработка новых типов конструкций устройств позиционирования, совместимых с современной аппаратурой радиолокационных средств миллиметрового диапазона длин волн, неизменно остается актуальной.

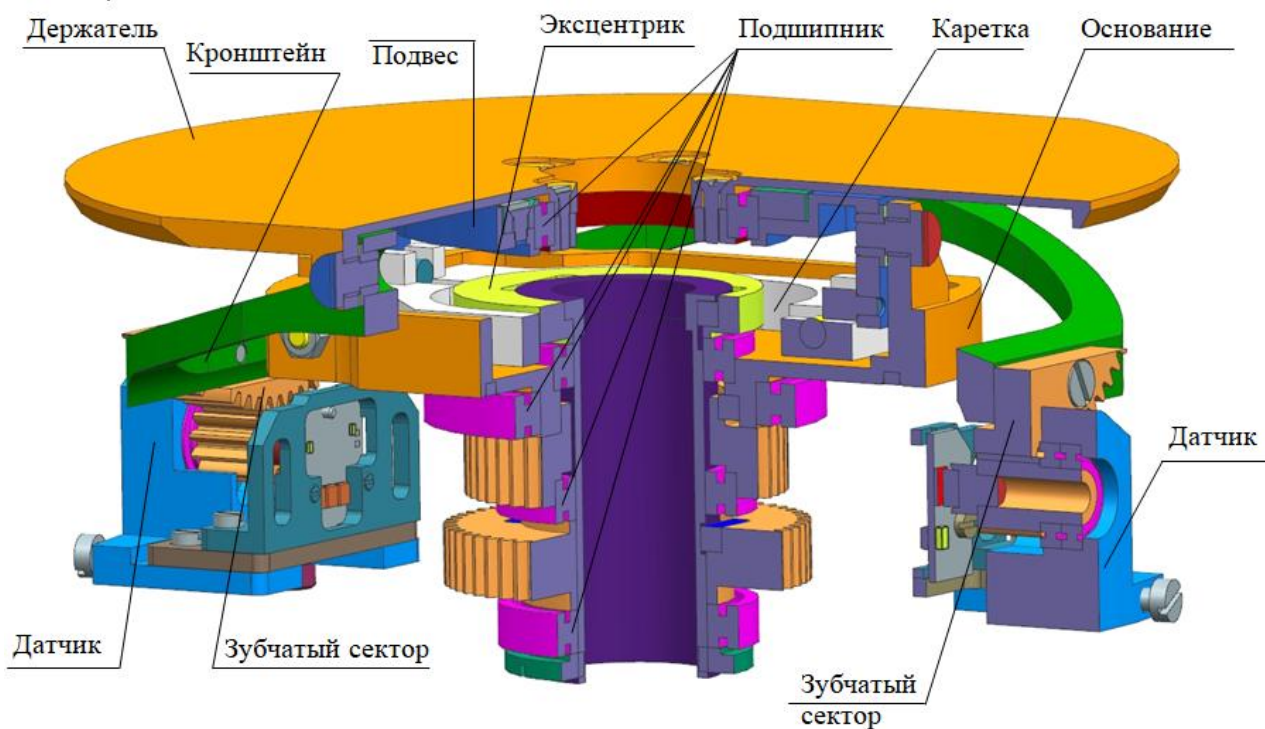
В представляемой работе описывается новая конструкция устройства углового позиционирования бортового радиолокатора миллиметрового диапазона длин волн с магнитометрической системой встроенного контроля положения луча. Конструкция устройства обеспечивает компоновочную совместимость с приемопередающим модулем, характеризуется компактностью и относительно малыми массогабаритными характеристиками. Примененные современные неконтактные магнитометрические методы измерения механического положения движущихся элементов конструкции позволяют повысить точность измерения угловых координат и надежность работы всей системы без увеличения размеров и массы, а также предохраняют от сопутствующих повреждений конструктивных элементов системы во время отработки алгоритмов управления и тестовых запусков.

### 1. Конструкция радиолокатора

Трехмерная модель разработанного устройства позиционирования (УП) представлена на рис.1, кинематическая схема – на рис.2. Устройство обеспечивает обзор области пространства в диапазоне углов азимута и элевации  $\pm 30^\circ$  лучом шириной  $1,5^\circ$  за 1,25 с (получено расчетным путем).

Устройство позиционирования разработано для антенны, построенной по схеме антенны с поворотом плоскости поляризации. Она состоит из неподвижного отражателя 3, установленного в радиопрозрачный кожух 2, подвижного отражателя 4, устанавливаемого на УП, и облучателя, находящегося в центре УП (на рисунках не показан).

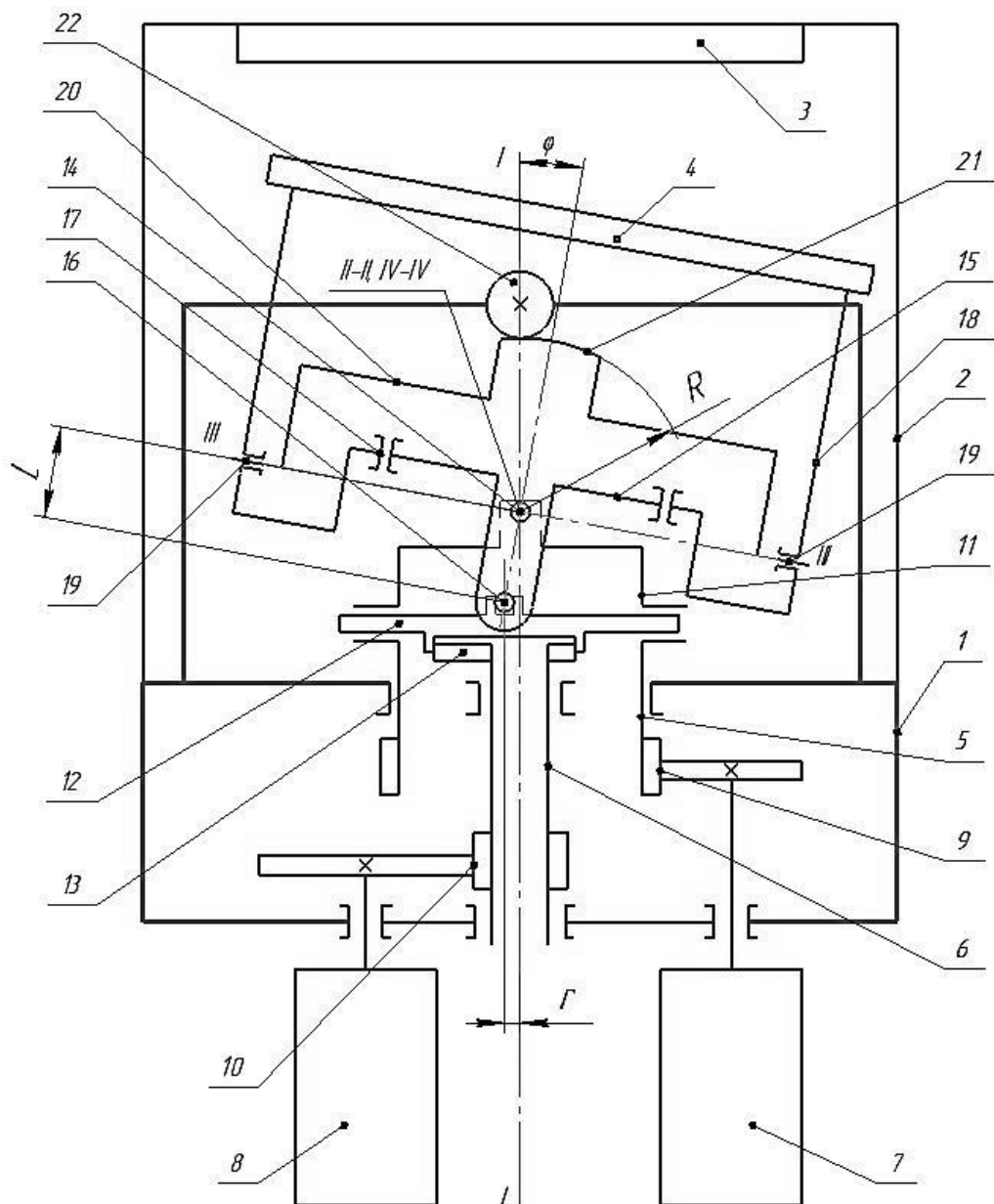
Механическое устройство позиционирования (рис.2) включает полые вал азимутального вращения 5 и вал угломестного вращения 6, расположенные соосно в корпусе 1, вращающиеся вокруг строительной оси устройства I-I и приводимые в движение двумя двигателями: вал 5 - двигателем 7, вал 6 - двигателем 8 через зубчатые передачи 9 и 10 соответственно. На оси полых валов азимутального и угломестного вращения установлен облучатель антенны (на рисунках не показан). Основание 11 жестко связано с валом азимутального вращения 5. Каретка 12 расположена на основании 11 с возможностью линейного перемещения относительно основания в плоскости, перпендикулярной строительной оси устройства. Эксцентрик 13, установленный в торце вала 6, при вращении относительно основания 11 приводит к возвратно-поступательным движениям каретки 12. На основании 11 посредством штырей 14 также установлен подвес 15, который под действием каретки 12 наклоняется вокруг оси закрепления II-II, перпендикулярной строительной оси устройства I-I, при этом ось II-II вращается вокруг оси I-I вместе с основанием. На подвесе 15 на расстоянии  $L$  от его оси крепления установлены штифты 16, которые находятся в пазах каретки 12. Таким образом, линейные перемещения каретки 12 приводят к наклону подвеса 15 относительно основания 11 (вокруг оси II-II). Максимальный угол наклона подвеса  $\varphi$  определяется расстоянием  $r$  между центрами эксцентрика 13 и расстоянием  $L$ .



**Рис.1 – Трехмерная модель разработанного УП**

При работе устройства основание 11 с подвесом 15 вращаются вокруг строительной оси устройства I-I. Для обеспечения возможности построения антенны с поворотом плоскости поляризации необходимо ограничить вращение подвижного отражателя вокруг своей оси. Для этого на подвес 15 через подшипник 17 установлен держатель 18, на котором посредством штырей 19 установлен кронштейн 20, способный свободно вращаться вокруг оси закрепления относительно держателя (вокруг оси III-III). Кронштейн 20 имеет зубчатые секторы 21, взаимодействующие с шестернями 22, установленными на корпусе 1. Центры зубчатых секторов 21 находятся на линии IV-IV, параллельной оси установки шестерней 22 и проходящей через точку пересечения строительной оси устройства I-I и оси II-II крепления подвеса 15 и основания 11. При работе устройства подвижный отражатель 4, установленный на держателе 18, будет наклоняться вокруг оси III-III относительно кронштейна 20, а также наклоняться вокруг оси IV-IV совместно с кронштейном 20, при этом ось IV-IV остается неподвижной относительно корпуса 1, а ось III-III вращается вокруг оси IV-IV в плоскости, перпендикулярной оси IV-IV и проходящей через точку пересечения осей I-I и IV-IV. На оси шестерней 22 установлены магниты системы встроенного контроля угломестного положения луча, построенной на основе магнитометрических датчиков угла вращения. Магниты системы встроенного контроля азимутального угла обзора конструктивно размещаются на основании, датчики магнитного поля – на неподвижном корпусе устройства позиционирования.





**Рис.2 – Кинематическая схема разработанного УП:**

1 – корпус, 2 – радиопрозрачный кожух, 3 – неподвижный отражатель, 4 – подвижный отражатель, 5 – вал азимутального вращения, 6 – вал угломестного вращения, 7, 8 – двигатели, 9, 10 – зубчатые передачи, 11 – основание, 12 – каретка, 13 – эксцентрик, 14 – штырь, 15 – подвес, 16 – штырь, 17 – подшипник, 18 – держатель, 19 – штырь, 20 – кронштейн, 21 – зубчатый сектор, 22 – шестерня.

## 2. Математическая модель УП

Математическая модель УП была построена для управления пространственным положением сканирующего луча радиолокатора. Она основана на кинематических зависимостях между конструктивными элементами радиолокатора. Задачей моделирования является определение зависимости угловых положений азимутального  $\alpha_{об.1}(t)$  и угломестного  $\alpha_{об.2}(t)$  двигателей от углов  $\alpha(t)$  и  $\varphi(t)$  в полярной системе координат, привязанной к радиолокатору, и от координат  $x(t)$ ,  $y(t)$  в декартовой системе, где  $x(t)$ ,  $y(t)$  – координаты точки пересечения оси сканирующего луча с картинной плоскостью (плоскостью, перпендикулярной строительной оси УП).



Эксцентрик при равномерном вращении относительно основания приводит к возвратно-поступательным движениям каретки по гармоническому закону (рис.1):

$$l(t) = r \sin(\Delta\alpha(t)), \quad (1)$$

где  $r$  – расстояние между центрами эксцентрика;

$\Delta\alpha(t) = [\alpha_1(t) - \alpha_2(t)]$  – угол поворота одного вала относительно другого;

$\alpha_1(t)$  – угол поворота вала азимутального вращения;

$\alpha_2(t)$  – угол поворота вала угломестного вращения.

Также на основании посредством штырей установлен подвес, который под действием каретки наклоняется вокруг оси закрепления II-II, перпендикулярной строительной оси устройства I-I, при этом ось II-II вращается вокруг оси I-I вместе с основанием. На подвесе на расстоянии  $L$  от его оси крепления установлены штифты, которые находятся в пазах каретки. Таким образом, линейные перемещения каретки приводят к наклону подвеса относительно основания (вокруг оси II-II) по следующему закону:

$$\varphi(t) = \arcsin \frac{r \sin(\Delta\alpha(t))}{L}. \quad (2)$$

Поворот УП по азимуту  $\alpha$  осуществляется синхронным вращением основания и эксцентрика, приводимых в движение двигателями 1 и 2 через зубчатые передачи с коэффициентами 1 и 43/57 соответственно. При этом эксцентрик не будет перемещать каретку, а подвес будет вращаться вокруг оси I-I без изменения значения угла  $\varphi$ . Таким образом, чтобы обеспечить поворот УП по азимуту по заданному закону  $\alpha(t)$ , с учетом коэффициентов передачи двигателя должны вращаться следующим образом:

$$\alpha_{об.1}(t) = \alpha(t); \quad (3)$$

$$\alpha_{об.2}(t) = \frac{57}{43} \alpha(t). \quad (4)$$

Поворот отражателя по углу места  $\varphi$  осуществляется поворотом эксцентрика относительно основания по закону (2). Чтобы найти законы изменения угловых положений двигателей по заданным  $\alpha(t)$  и  $\varphi(t)$ , необходимо в формулу (4) добавить дополнительное слагаемое из формулы (2):

$$\alpha_{об.1}(t) = \alpha(t); \quad (5)$$

$$\alpha_{об.2}(t) = \frac{57}{43} \left( \alpha(t) + \arcsin \frac{L \sin(\varphi(t))}{r} \right). \quad (6)$$

Зависимость  $\alpha(t)$  и  $\varphi(t)$  от задаваемых  $x(t)$  и  $y(t)$  определяется в виде

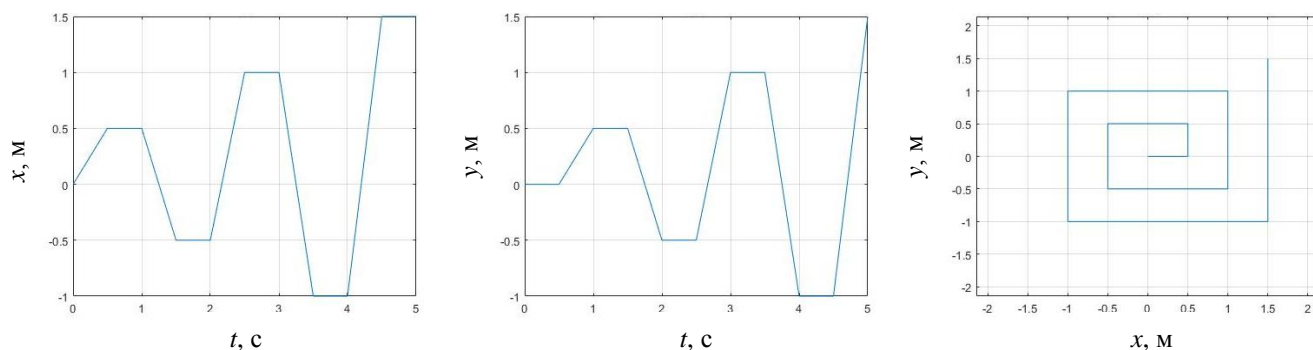
$$\alpha(t) = \arctg \frac{y(t)}{x(t)}; \quad (7)$$

$$\varphi(t) = \arctg \frac{\sqrt{(x(t))^2 + (y(t))^2}}{r_z}. \quad (8)$$

где  $r_z$  – расстояние от радиолокатора до картинной плоскости.

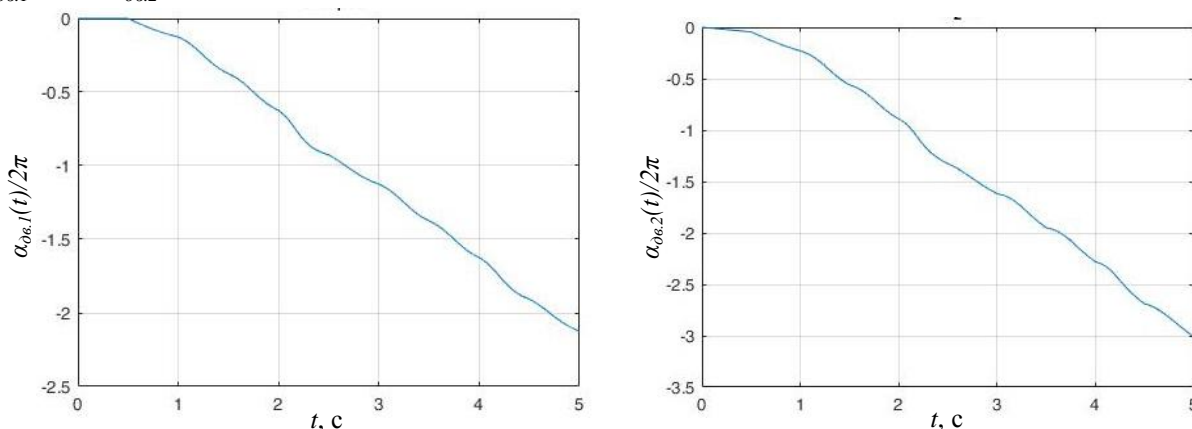
Если полученные равенства (7), (8) подставить в равенства (5), (6), то получим закон изменения углов поворота двигателей в зависимости от задаваемого закона сканирования в виде  $x(t)$  и  $y(t)$ .

С помощью описываемой конструкции можно реализовать любой закон сканирования в пределах конструктивных ограничений, определяемых размерами  $L$  и  $r$ . Для этого необходимо задать желаемый закон сканирования в виде  $\alpha(t)$  и  $\varphi(t)$  или в виде  $x(t)$  и  $y(t)$ . В качестве примера рассмотрим закон сканирования, представленный на рис.3.



**Рис.3 – Закон сканирования, заданный в картинной плоскости**

Далее по формулам (5), (6) находим законы изменения углов поворота двигателей  $\alpha_{де.1}(t)$  и  $\alpha_{де.2}(t)$ , представленные на рис.4.



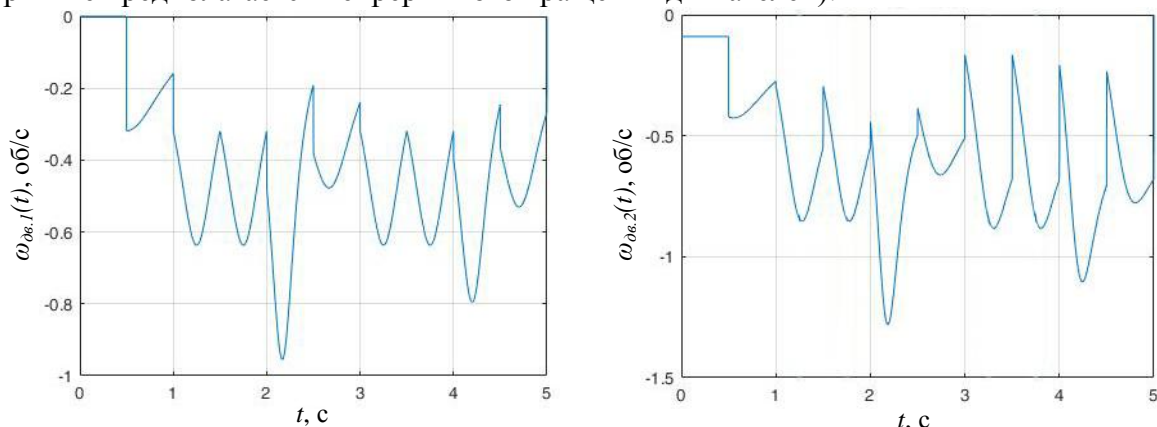
**Рис.4 – Графики изменения углов поворота двигателей  $\alpha_{де.1}(t)$  и  $\alpha_{де.2}(t)$**

### 3. Анализ законов сканирования радиолокатора

Наиболее распространенными законами сканирования являются: растровый (или линейный построчный) и спиральный [4]. Механическое устройство позиционирования будет являться ограничивающим фактором в скорости работы радиолокатора. Это связано с высокой инерционностью механической системы, обусловленной постоянными времени, на несколько порядков отличающимися от электрических. Например, для реализации закона сканирования, представленного на рис.3, необходимо, чтобы двигатели вращались с угловыми скоростями, представленными на рис.5. Резкие скачки расчетных значений угловых скоростей двигателей свидетельствует о невозможности реализации рассматриваемого закона сканирования ввиду инерционности механической системы и ограниченных динамических возможностей двигателей. Из этого следует, что закон сканирования нужно выбирать таким образом, чтобы режимы работы двигателей, с учетом инерционности и триботехнических потерь конструкции (в данной статье эти вопросы не рассматриваются), не превышали предельно допустимых нагрузок.

В представленной конструкции двигатели могут совершать неограниченное количество оборотов. При этом если валы азимутального и угломестного вращения будут двигаться с постоянными, но разными скоростями ( $\Delta\alpha \neq 0$ ), азимут  $\alpha(t)$  будет линейно возрастать (согласно формуле (5)), а угол места  $\varphi(t)$  будет меняться по гармоническому закону (2). Точка пересечения оси сканирующего луча с картинной плоскостью будет перемещаться по спиральному закону (для антенн с поворотом плоскости поляризации угол наклона оси сканирующего луча в два раза превышает угол наклона нормали твистрефлектора). Такой закон сканирования будет оптимальным для данной конструкции, поскольку при работе радиолокатора необходимо поддерживать постоянные скорости вращения электродвигателей без необходимости их периодического разгона и торможения (в

отличие от конструкций, построенных на основе механизмов вращения, описанных в [3], в которых не предполагается непрерывного вращения двигателей).



**Рис.5 – Графики изменения угловых скоростей двигателей  $\omega_{об.1}(t)$  и  $\omega_{об.2}(t)$**

Далее проведем теоретическую оценку минимально возможного времени получения полного кадра при условии, что максимальная скорость вращения двигателей составляет 3000 об/мин. Временем получения полного кадра ( $T_K$ ) будем считать время, необходимое УП для перемещения нормали твистрефлектора по углу места от нулевого положения (совпадающего со строительной осью изделия) до амплитудного значения, определяемого конструкцией (размерами  $r$  и  $L$  на рис.2) при условии, что минимальное угловое расстояние между соседними витками спирали будет равно ширине диаграммы направленности сканирующего луча. Согласно данному определению, за время одного оборота эксцентрика относительно основания проходит 4 кадра:

$$v_{экс} = v_{осн} \pm \frac{1}{4T_K}. \quad (9)$$

С учетом коэффициентов зубчатых передач получаем:

$$v_1 = \frac{43}{57} v_2 \pm \frac{1}{4T_K}. \quad (10)$$

Если максимальная скорость вращения двигателей ограничена значением в 3000 об/мин (50 об/с), из равенства (10) получаем, что скорость вращения двигателя 1 не должна превышать значения 37,72 об/с. Для упрощения дальнейших расчетов, будем считать, что двигатель 1 (и соответственно вал азимутального вращения) будет вращаться со скоростью 32 об/с.

Количество витков в кадре будем рассчитывать из условия того, что за первый полный оборот по азимуту (в момент времени  $T_a$ ), угол  $\varphi$  должен измениться на ширину ДН (первым считается оборот, когда угол  $\varphi$  возрастает от 0, так как в этот момент скорость изменения угла  $\varphi$  максимальна). Согласно формуле (2) имеем:

$$\varphi(T_a) = 2 \arcsin \frac{r \sin(\Delta\alpha(T_a))}{L}. \quad (11)$$

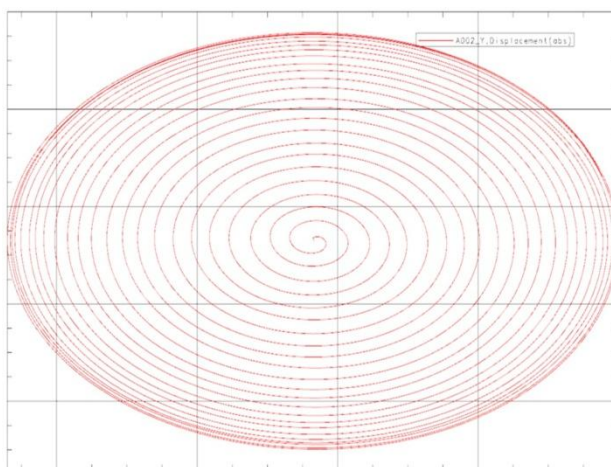
Из формулы (11) с учетом равенства (9) получаем:

$$T_K = \frac{\pi \cdot T_a}{2 \arcsin \left( \frac{L}{r} \sin \frac{\varphi(T_a)}{2} \right)}. \quad (12)$$

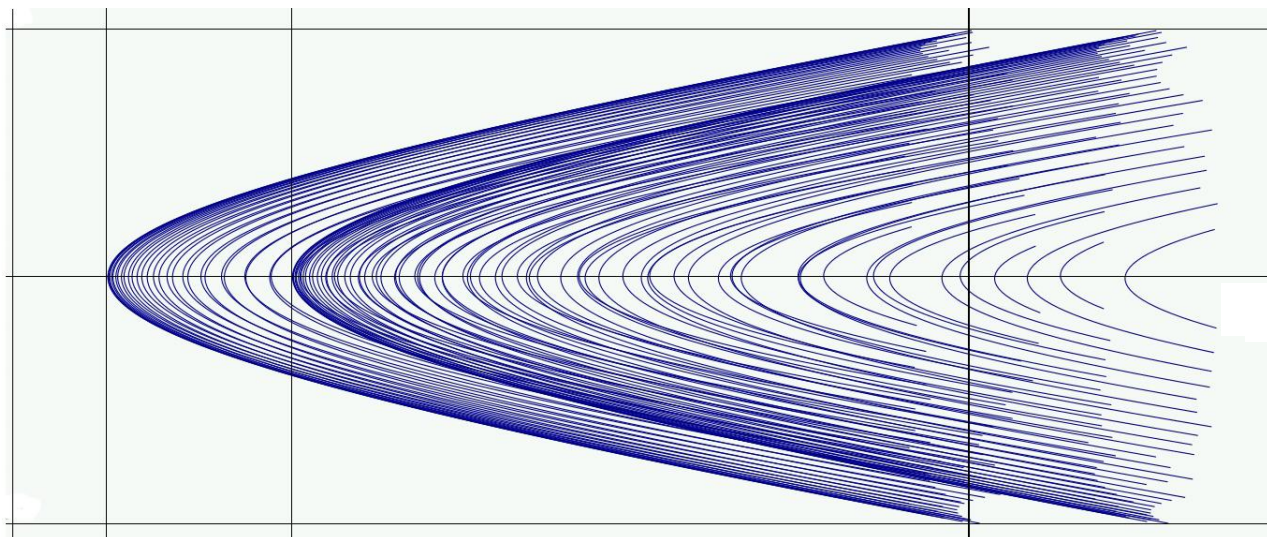
Подставляя значения  $T_a=1/32$  с,  $\varphi(T_a)=1,5^\circ$  (ширина ДН), получаем  $T_K \approx 1,2$  с. Количество витков в спирали будет равно  $N=T_K/T_a \approx 39,1$ . Для упрощения примем  $N=40$ , тогда минимально возможное время получения полного кадра для данной конструкции будет

равно  $T_k=40/32=1,25$  с. На рис.6 представлена линия пересечения нормали твистрефлектора с картинной плоскостью за время, равное времени одного кадра.

В качестве примера рассмотрим случай использования разрабатываемого УП на бортовом радиолокаторе для поиска наземных целей. На рис.8 представлены линии пересечения оси сканирующего луча с земной поверхностью при горизонтальном полете (строительная ось РЛ параллельна подстилающей поверхности) на высоте 2000 м с постоянной скоростью за время, равное двум кадрам. Как видно из рис.8, часть времени, за которое происходит сканирование полного кадра, нормаль твистрефлектора не попадает на сканируемую поверхность. Однако, благодаря высокой скорости вращения УП относительно скорости перемещения самого РЛ, можно добиться такого сочетания задаваемых характеристик УП, при котором обзор сканируемого пространства на каждом кадре будет происходить без пропусков. При прямолинейном движении РЛ большая часть сканируемой области обозревается несколько раз (за счет наложения кадров, как показано на рис.7). Это увеличивает вероятность нахождения цели в условиях сложного рельефа за счет того, что одна и та же локальная область сканируемого пространства будет просматриваться под немного разными углами.



**Рис.6 – Линия пересечения нормали твистрефлектора с картинной плоскостью**



**Рис.7 – Линии пересечения нормали твистрефлектора со сканируемой поверхностью**

#### **4. Магнитометрическая система автоматического контроля**

Поскольку для приведения в движение антенной системы радиолокатора применяются сложные кинематические схемы с большим числом звеньев, для повышения надежности системы без изменения массогабаритных характеристик в её состав включены

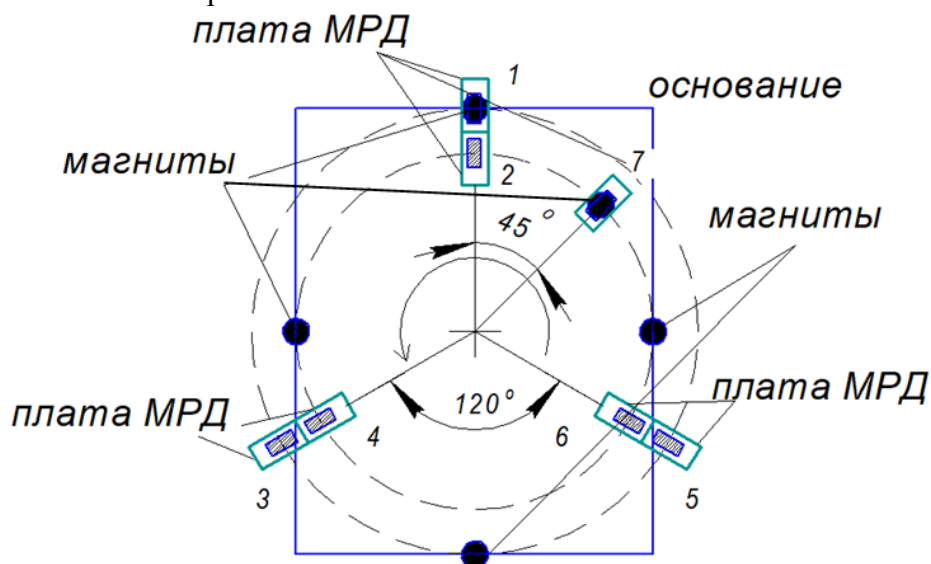
элементы автоматического встроенного контроля положения. Магнитометрическая система встроенного контроля углового положения радиолокатора включает в себя угловые датчики вращения кронштейна и систему датчиков присутствия магнитов для секторального отслеживания азимутального угла сканирования луча и фиксации нулевого отсчета.

Элементы магнитометрического контроля размещены как можно ближе к конструктивным деталям УП, ответственным за позиционирование радиолокатора. Это обеспечивает возможность апостериорной коррекции ошибок позиционирования, связанных с конструктивными допусками и погрешностями изготовления деталей механизма.

Для отслеживания угломестных координат положения луча может быть использована отечественная микросхема датчика углового положения диаметрально намагниченного торцевого магнита 1389НМ025 от ЗАО «НТЦ Зеленоградский нанотехнологический центр» [5]. Она производит вычисление кода положения магнита, расположенного на торце вала, связанного с шестерней, находящейся в кинематической зависимости от вращательных перемещений кронштейна. Для точного позиционирования микросхемы в составе прибора применены ряд конструктивных настроечных элементов.

Для отслеживания азимутального положения луча радиолокатора в секторальных областях  $\Delta\varphi=30^\circ$  может быть использована следующая мехатронная схема (рис.10). Система основана на применении магниторезистивных пороговых чувствительных элементов производства филиала РФЯЦ-ВНИИЭФ «НИИИС им. Ю.Е. Седакова» (1 мост Уинстона) [6]. Магниторезистивные пленки не чувствительны к полярности и нормальной компоненте магнитного поля и обеспечивают максимальный уровень сигнала при нулевом угле направления напряженности магнитного поля.

Сдвоенные платы магниторезистивных датчиков (МРД) устанавливаются под углом  $120^\circ$  по отношению друг к другу на корпус. Для фиксации нулевого положения вводится еще один датчик под углом в  $45^\circ$  к первому. По два магнита устанавливаются в серединах длинных и коротких стенок вращающегося основания, дополнительный магнит размещается под углом  $45^\circ$ , как показано на рис.8.



**Рис.8 – Схема секторального контроля азимутального угла**

Рабочие области датчиков размещаются на окружностях вращения магнитов. В случае непрерывного вращения основания в моменты времени, соответствующие дискретным отсчетам азимутального угла в  $30^\circ$ , будут наблюдаться сигналы от магниторезистивных пороговых датчиков в следующем порядке: 1, 4, 5, 2+7, 3, 6, 1, 4, 5, 2, 3, 6, начиная с положения, изображенного на рис.10 при вращении по часовой стрелке. При смене направления вращения двигателей максимальный угловой дискрет диагностики



азимутального угла составит  $\Delta\varphi=210^\circ$ , минимально возможный –  $15^\circ$ . Нулевое положение фиксируется по одновременному наличию сигналов со 2-го и 7-го датчиков.

### Выводы

Результатом разработки стало создание конструкции, обеспечивающей механическое высокоскоростное сканирование области цели по спиральному закону с возможностью управления траекторией движения сканирующего луча в процессе работы. Введение дополнительных элементов конструкции позволило использовать антенну с поворотом плоскости поляризации. Предложен способ секторального азимутального отслеживания текущего положения луча с фиксацией нулевого отсчета и косвенный метод контроля текущего угла элевации, реализованные на отечественной элементной базе.

### Библиографический список

1. **Нечаев, Е.Е., Дербин, К.С.** Современные бортовые радиолокационные станции и антенные решетки многофункциональных авиационных комплексов военного назначения / Научный вестник МГТУ ГА. №221, 2015, с.90-105.
2. Метеолокаторы / [www.airteam.eu/ru/p/garmin-gwx-70-3](http://www.airteam.eu/ru/p/garmin-gwx-70-3) (дата обращения 29.03.2022).
3. **Ривкин, С.С.** Стабилизация измерительных устройств на качающемся основании. - М. «Наука», 1978г., стр. 197-198.
4. **Никольский, Б.А.** Бортовые радиоэлектронные системы [Электронный ресурс]: учеб. / Б.А. Никольский. - Самара: Изд-во Самар., гос. аэрокосм. ун-та, 2013. - 1 электр. опт. диск.
5. Микросхема датчика углового положения диаметрально намагниченного торцевого магнита K1382NM025\_datasheet.pdf / [www.zntc.ru](http://www.zntc.ru) (дата обращения 29.03.2022).
6. Магниторезистивные чувствительные элементы для пороговых датчиков / <https://www.niis.nnov.ru/activity/ntb/gm> (дата обращения 29.03.2022).

## ИМИТАЦИОННАЯ МОДЕЛЬ РАДИОЛОКАЦИОННОГО СИГНАЛА, ОТРАЖЕННОГО ОТ ШЕРОХОВАТОЙ ПОДСТИЛАЮЩЕЙ ПОВЕРХНОСТИ

А.В. Анфимова

ORCID: 0000-0003-4510-6880 e-mail: aav60892@mail.ru

М.С. Дубровин

ORCID: 0000-0002-0351-6104 e-mail: mx325999@gmail.ru

Ю.М. Куликов

ORCID: 0000-0002-8549-8906 e-mail: umkulikov@yandex.ru

А.В. Назаров

ORCID: 0000-0002-0261-1724 e-mail: nazarov52@mail.ru

Филиал ФГУП РФЯЦ-ВНИИЭФ «НИИИС им. Ю.Е. Седакова»

Доклад посвящен моделированию огибающей сигнала, отраженного от шероховатой подстилающей поверхности, с использованием фасетной модели. Математические модели отраженных от подстилающей поверхности радиолокационных сигналов необходимы при синтезе и анализе работы различных радиолокационных систем. Целью исследования является разработка программной модели, обеспечивающей корректный расчет энергетических характеристик и корреляционных свойств принимаемых радиолокационных сигналов и последовательностей различной длительности. Отличительной чертой реализованной математической модели является возможность описания не только изменения формы огибающей сигнала при отражении от подстилающей поверхности, но и ее изменения во времени, возникающего при наличии взаимного движения радиолокационной станции и фасетов, без использования априорной информации о статистическом законе флуктуаций. Показана возможность применения модели для расчетов сигналов, отраженных от сложных радиолокационных сцен, содержащих объекты сложной геометрии на фоне подстилающей поверхности.

**Ключевые слова:** радиолокация, фасетная модель, огибающая сигнала, флуктуации отраженного сигнала.

## SIMULATION MODEL OF THE RADAR BACKSCATTER FROM A ROUGH SURFACES

A.V.Anfimova, M.S.Dubrovina, Y.M.Kulikov, A.V.Nazarov

Branch of federal state unitary enterprise «Russian federal nuclear center - All-russian scientific-research institute of experimental physics» «Scientific-research institute of measuring systems n.a. Yu.E. Sedakov»

**Purpose:** The purpose of research is to develop a software model of the envelope signal reflected from a rough surface and provide correct energy characteristics and correlation properties of received radar signals and sequences of various durations.

**Design/methodology/approach:** The paper deals with the modeling of the envelope signal reflected from a rough surface using a facet model. The envelope signal calculations were made by using CUDA technology and parallel calculation methods. Signal propagation model considers the laws of geometric optics and ray-tracing algorithm.

**Findings:** Original model simulates changes of the signal envelope amplitude upon reflection from the surface, also its changes in time, which occurs in the presence of mutual motion of the radar station and the surface, without using a priori information about the statistical law of fluctuations.

**Research limitations/implications:** Mathematical models of radar signals reflected from the surface are necessary for the designing and analysis of operations of various radar systems.

**Originality/value:** The paper shows the fluctuations simulation using phenomenological method. Original model calculates envelope signal reflected from complex radar scenes including objects of complex geometry and background surface.

**Keywords:** radar, facet model, signal envelope, amplitude fluctuation.

Задача расчета огибающей радиолокационных сигналов, отраженных от подстилающей поверхности, возникает при разработке различных радиолокационных систем и решается с целью определения требуемых параметров как самих систем, так и параметров используемых зондирующих сигналов. Существующая в настоящее время теоретическая база [1] дает возможность создавать имитационные модели, позволяющие воспроизводить различные характеристики принимаемых радиолокационной системой сигналов с высокой степенью достоверности [1,2]. При математическом моделировании процессов распространения и отражения зондирующих сигналов наиболее строгим является подход, который заключается в решении электродинамической задачи распространения электромагнитных волн и их рассеяния на подстилающей поверхности [3,4]. При данном подходе подстилающая поверхность задается геометрическими характеристиками и электромагнитными параметрами (проводимостью, диэлектрической и магнитной проницаемостью). Применение данного подхода связано со значительными математическими трудностями, обусловленными сложностью описания реальных подстилающих поверхностей и очень большим объемом вычислений. Кроме того, на практике чаще всего известны лишь интегральные характеристики подстилающей поверхности: коэффициент отражения, диаграммы обратного рассеяния и т.д. Рельеф подстилающей поверхности описывается перепадом высот, максимальными уклонами, радиусом корреляции, что также затрудняет использование электродинамической модели рассеяния сигналов на протяженной поверхности.

Другой подход к моделированию заключается в использовании феноменологического подхода - концепции независимых отражателей. В этом случае подстилающая поверхность представляется в виде совокупности независимых отражателей, а отраженный сигнал - суперпозицией элементарных отраженных сигналов. Представление поверхности в виде элементарных отражателей, несмотря на ряд используемых допущений (независимость элементарных отражателей, равномерный закон распределения фаз), широко применяется при моделировании радиолокационных устройств, работающих по протяженным объектам [4].

В настоящей работе для построения модели отраженного от подстилающей поверхности сигнала используется феноменологический подход. Как было отмечено выше, согласно феноменологическому подходу, протяженный радиолокационный объект или подстилающую поверхность можно представить в виде набора дискретных отражателей – фацетов[4]. Тогда сигнал на входе приемной антенны равен сумме сигналов, отраженных от всех видимых статистически независимых фацетов. Этот принцип позволяет для каждого из фацетов заменить сложные пространственно-электрические характеристики реальных поверхностей на обобщенную характеристику – эффективную поверхность рассеяния, заданную для фиксированных поляризации, длины волны и угла облучения, зависящую от типа подстилающей поверхности – диаграмму обратного рассеивания (ДОР).

Для корректной обработки приборов последовательности сигналов, полученные путем моделирования, должны повторять не только энергетические свойства реальных сигналов, но и статистические. Мгновенные отсчеты сигналов, полученные в процессе моделирования, должны иметь корреляционные связи, соответствующие корреляционным свойствам реальных сигналов. Наиболее простым с точки зрения практической реализации является метод моделирования, суть которого заключается в расчете средней формы отраженного сигнала и моделирования флуктуаций путем использования известного статистического закона распределения флуктуаций по амплитуде и априорной информации об их корреляционных свойствах [5,6]. Недостатком данного подхода является невозможность учета пространственной корреляции принимаемых сигналов. Для корректной обработки приборов, использующих информацию о пространственной корреляции сигналов, ее учет в модели является принципиально важным.

В настоящей работе приведено описание математической модели, использующей подход, в рамках которого подстилающая поверхность представляется в виде набора независимых дискретных отражателей со случайной составляющей коэффициента отражения  $A_i$ , обусловленной неидентичностью каждого из фацетов, и случайной составляющей фазы  $\varphi_i$ , возникающей вследствие её скачка при отражении. Теоретические основы данного подхода изложены в работах [1, 4]. Учет  $A_i$  и  $\varphi_i$  является принципиальным отличием от моделей



[5,6,7]и позволяет корректно учесть пространственную корреляцию принимаемых сигналов без использования априорной информации о корреляционных свойствах реальных сигналов. Случайная амплитуда в рассматриваемой модели описывается рэлеевским распределением, случайная фаза равномерно распределена на интервале от 0 до  $2\pi$ .

Мощность отражённого сигнала от каждого facets  $P_i$  рассчитывается по основному уравнению радиолокации:

$$P_i(t) = \frac{P_{\text{изл}}(t-\tau)G(\varphi,\vartheta)^2\sigma_i\lambda^2}{(4\pi)^3R_i^4}, \quad (1)$$

где  $P_i(t)$  и  $P_{\text{изл}}(t)$  – зависимость мощности сигнала, принятого приемником и излученного передатчиком, от времени соответственно;

$\tau$  – задержка распространения сигнала, равная  $\tau=2R_i/c$ ;

$G(\varphi,\vartheta)$  – коэффициент усиления антенны в направлении на facets;

$\sigma_i$  – ЭПР facets, рассчитываемая по формуле (2);

$\lambda$  – длина волны;

$R_i$  – расстояние между фазовым центром АФС и центром facets.

ЭПР facets  $\sigma_i$  рассчитывается по формуле:

$$\sigma_i = \sigma_0(\theta) \cdot S_i, \quad (2)$$

где  $\sigma_0(\theta)$  – удельная ЭПР facets, зависящая от угла облучения (диаграмма обратного рассеяния),  $S_i$  – площадь facets.

Для получения средней формы мощности сигнала  $P_\Sigma(t)$ , отраженного от всей подстилающей поверхности, выполняется суммирование отраженных сигналов от всех  $N$  facets, для которых выполняется условие прямой видимости:

$$P_\Sigma(t) = \sum_{i=1}^N P_i(t), \quad (3)$$

Сумма элементарных сигналов с учетом времени их распространения до точки приема образует среднюю форму огибающей отраженного сигнала, которая не учитывает явлений интерференции и явлений, происходящих в следствии движения приемопередатчика. Мгновенная форма отраженного сигнала даже от ровной однородной поверхности в значительной степени отличается от средней и подвержена интенсивным флуктуациям [4] вследствие шероховатости реальных поверхностей. В общем случае отраженный от земной поверхности сигнал подвержен двум видам флуктуаций: межимпульсной («медленные флуктуации») и внутриимпульсной («быстрые флуктуации») [5].

В рамках модели отраженный сигнал представляет собой сумму парциальных сигналов, рассеянных элементами поверхности, каждый из которых является рэлеевским узкополосным процессом (ширина спектра мала по сравнению со средней частотой этого спектра) и может быть представлен в виде суммы двух некоррелированных квадратурных компонент. Каждая из этих компонент является нормальным процессом с нулевым средним [1,9,10].

Вычисление квадратурных составляющих случайного сигнала можно производить путем суммирования элементарных сигналов по текущей облучаемой площади с учетом фазы отраженного сигнала, зависящей от расстояния до отражателя, а также фазы и амплитуды случайного коэффициента отражения и мощности отраженной от каждого facets, рассчитываемой по (1):

$$U_c(t) = \sum_{i=1}^N \sqrt{P_i(t)} A_i \cos\left(\frac{4\pi}{\lambda} R_i + \varphi_i\right), \quad (4)$$

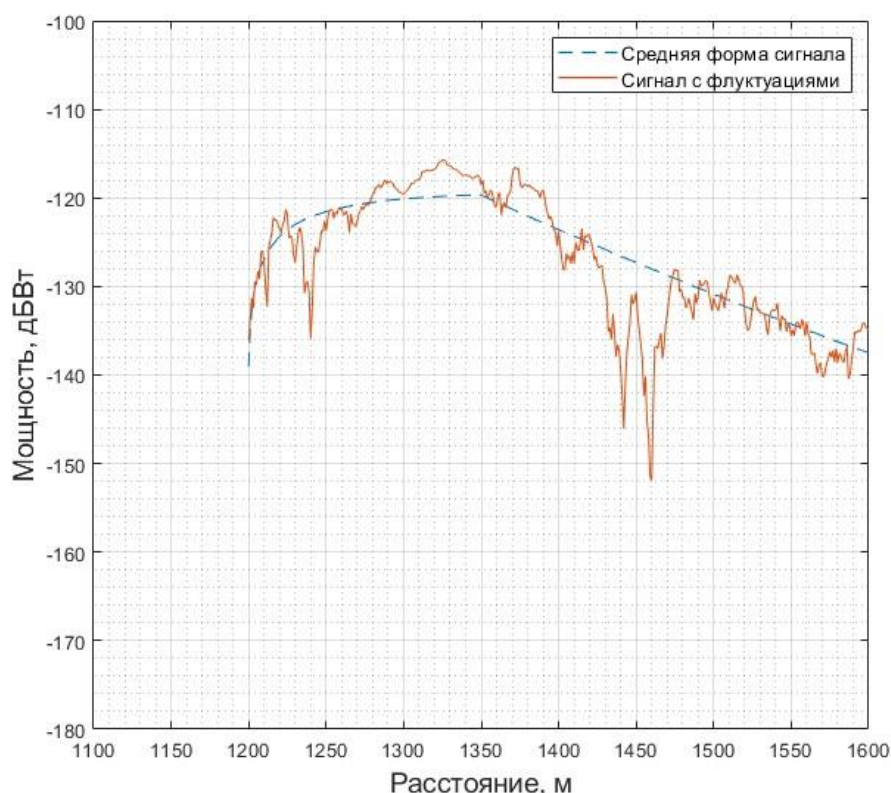
$$U_s(t) = \sum_{i=1}^N \sqrt{P_i(t)} A_i \sin\left(\frac{4\pi}{\lambda} R_i + \varphi_i\right), \quad (5)$$

где  $U_c$ ,  $U_s$  — косинусная и синусная составляющие отраженного сигнала;  $P_i(t)$  — мощность отраженного сигнала от  $i$ -го facets;  $R_i$  — расстояние до отражателя относительно точки излучения зондирующего сигнала;  $A_i$  — случайная составляющая коэффициента отражения  $i$ -го facets;  $\varphi_i$  — случайная составляющая фазы сигнала, отраженного от  $i$ -го facets; Мощность принятого сигнала в зависимости от времени может быть описана следующим выражением:

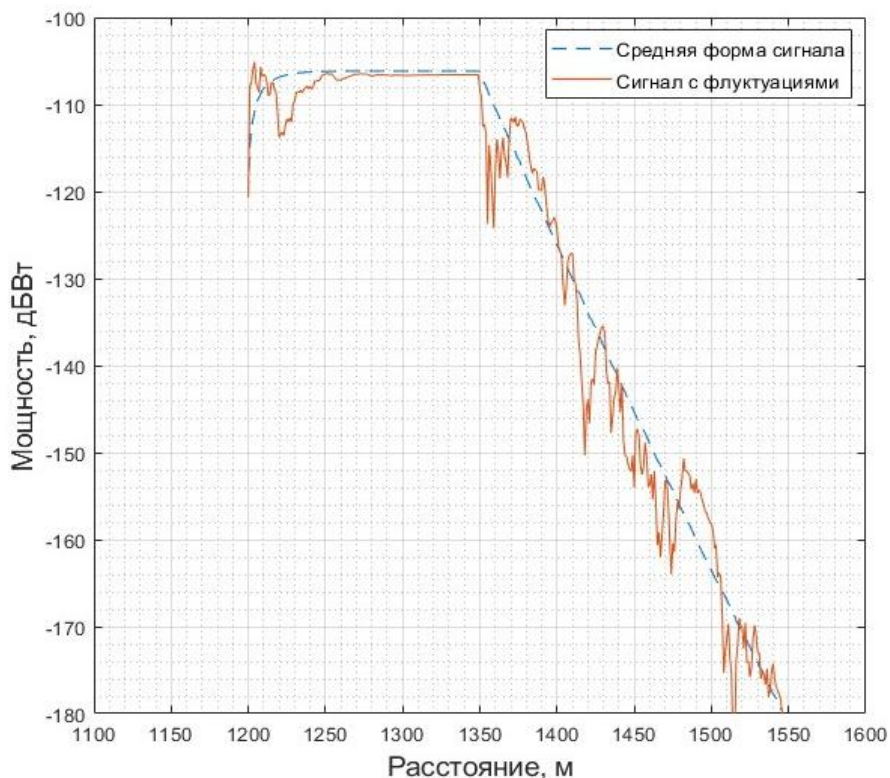
$$P(t) = U_c(t)^2 + U_s(t)^2. \quad (6)$$

На рис.1 и рис.2 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 над подстилающей поверхностью — 1200 м, длительность зондирующего импульса — 1 мкс, период следования зондирующих импульсов — 10 мкс, мощность передатчика — 1 Вт, ширина диаграммы направленности антенны —  $60^\circ$ , максимум диаграммы обратного рассеяния (ДНА) направлен в надир, скорость движения приемопередатчика — 280 м/с. Пунктирной линией на графиках представлена средняя форма огибающей принятого сигнала, сплошной — одна из реализаций формы мгновенного сигнала.

В описываемой модели алгоритм формирования флуктуаций и корреляционных зависимостей отсчетов принимаемых отраженных сигналов аналогичен причинам их возникновения в реальной ситуации.



**Рис. 1** Результат моделирования отраженного сигнала от поверхности типа «Заснеженный лес»



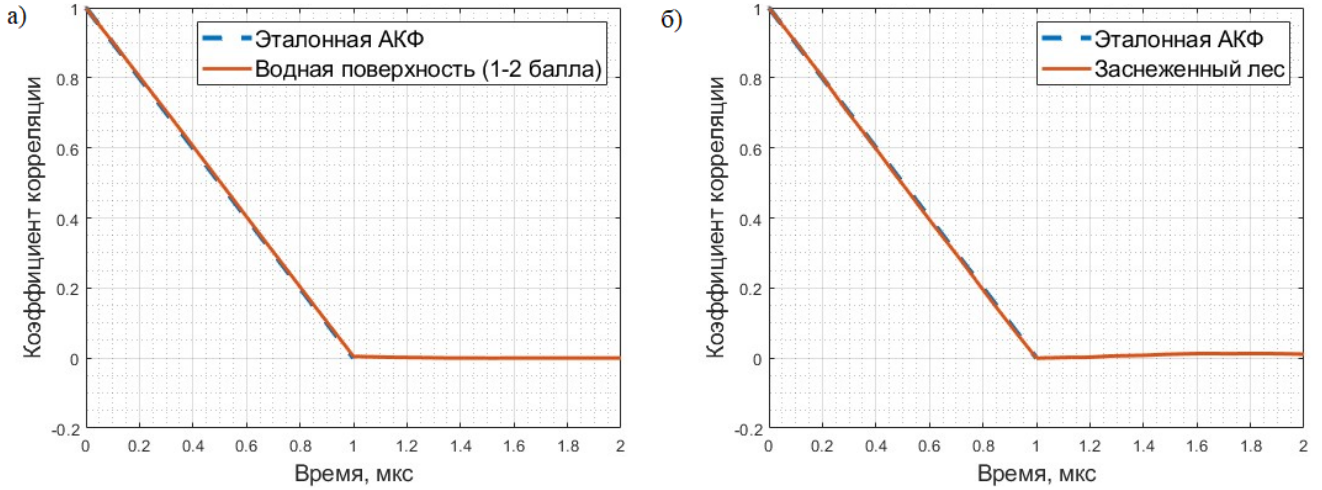
**Рис. 2** Результат моделирования отраженного сигнала от поверхности типа «Водная поверхность с волнением 1-2 балла»

Искажение формы мгновенных сигналов является следствием интерференции множества элементарных сигналов, отраженных от различных областей подстилающей поверхности. Амплитуда отраженных от земной поверхности сигналов, принимаемых движущимся радиолокатором, подвержена интенсивным флуктуациям вследствие изменения сдвига фаз сигналов, отраженных от различных участков облучаемой поверхности, в процессе движения. Время, соответствующее исчезновению корреляции выборок (время корреляции), обратно пропорционально ширине доплеровского спектра принимаемого сигнала. В формировании принимаемого отраженного сигнала в разные моменты времени участвуют различные участки поверхности. Когда интервал между моментами времени не превышает длительности излучаемого импульса, удается выявить общую часть поверхности, которая участвует в формировании отраженного сигнала в разные моменты времени, что и обуславливает наличие взаимной зависимости отсчетов сигналов.

С целью подтверждения корректности моделирования корреляционных связей было выполнено сравнение результатов моделирования с известными теоретическими результатами.

Корреляция отсчетов огибающих мгновенных реализаций сигналов определяется формой используемого зондирующего сигнала [5]. Для случая прямоугольного видеоимпульса форма внутриимпульсной автокорреляционной функции (АКФ) будет стремиться к треугольной (рис. 3).

### Внутриимпульсная корреляция



**Рис. 3 Усредненные АКФ:**

- а) сравнение эталонной АКФ и АКФ водной поверхности  
 б) сравнение эталонной АКФ и АКФ заснеженного леса

Анализ межимпульсной корреляции выполнялся путем вычисления АКФ значений импульсного среза для последовательности импульсов с дальнейшим усреднением. Были проанализированы корреляционные свойства для случаев подстилающей поверхности «водная поверхность» и «заснеженный лес».

Эталонные АКФ рассчитывались методом численного интегрирования с использованием выражения для расчета коэффициента взаимной корреляции для случая вертикального излучения с использованием отдельной передающей и двух приемных антенн. Коэффициент корреляции двух принимаемых сигналов в этом случае будет равен [9]:

$$\rho_{V_1 V_2}(t, t + \tau) = e^{j\omega_0 \tau} \frac{\iint \sqrt{A} \frac{\sqrt{B}}{R_2 R_1} e^{-j\frac{\omega_0}{c}(\Delta R + R_2 - R_1)} e^{j\left[\Phi\left(t' - \frac{R' + R_2}{c}\right) - \Phi\left(t - \frac{R + R_1}{c}\right)\right]} d\psi d\varphi}{\sqrt{\iint C \cdot R_1^{-2} d\psi d\varphi} \sqrt{\iint D \cdot R_2^{-2} d\psi d\varphi}}, \quad (7)$$

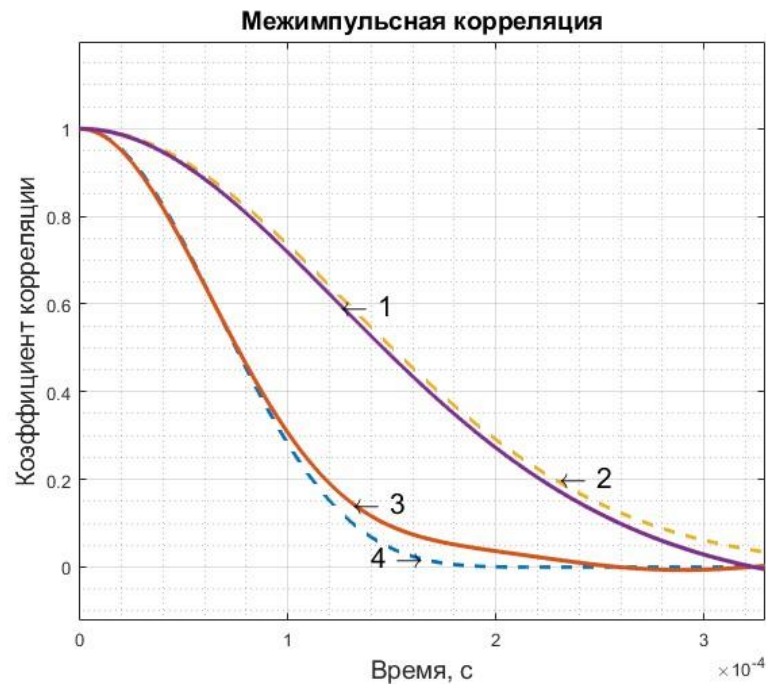
где  $A = G(\varphi, \psi)G(\varphi', \psi')G_1(\varphi_1, \psi_1)G_2(\varphi_2, \psi_2)$ ;  $B = \sigma(\varphi, \psi, \varphi_1, \psi_1)\sigma(\varphi', \psi', \varphi_1, \psi_1)$ ;  $C = G(\varphi, \psi)G_1(\varphi_1, \psi_1)\sigma(\varphi, \psi, \varphi_1, \psi_1)$ ;  $D = G(\varphi, \psi)G_2(\varphi_2, \psi_2)\sigma(\varphi, \psi, \varphi_2, \psi_2)$ ;  $\omega_0$  – средняя частота излучаемых непрерывных колебаний;  $\Phi(t)$  – функция, описывающая частотную и фазовую модуляцию сигнала, и являющаяся медленной функцией времени по сравнению с  $e^{-j\omega_0 \tau}$ ;  $\varphi, \psi$  – углы, определяющие положение отражателя относительно диаграммы направленности передающей антенн в двух взаимно перпендикулярных плоскостях, проходящих через ось диаграммы;  $\varphi_1, \psi_1$  – углы, определяющие положение отражателя относительно диаграммы направленности первой приемной антенны;  $G(\varphi, \psi)G_1(\varphi_1, \psi_1)$  – двумерные диаграммы направленности (по мощности) передающей и первой приемной антенны;  $\sigma$  – удельная эффективная площадь рассеяния;  $R_1$  – расстояние от отражателя до первой приемной антенны;  $R$  – расстояние от отражателя до передающей антенны;  $c$  – скорость распространения электромагнитных колебаний;  $t' = t + \tau$ ;  $\varphi', \psi', R'$  – значения  $\varphi, \psi, R$  в момент времени  $t'$ ;  $\Delta R = R' - R$ ; индексы «2» соответствуют второй антенне.

Переход к автокорреляционной функции огибающей сигнала для случая совмещенной приемной и передающей антенн в отсутствии фазовой модуляции можно выполнить путем замены индексов 2 на 1 и характеристик передающей антенны на параметры приемной.



$$\rho_{V_1 V_1}(t, t + \tau) = \frac{\iint G(\varphi, \psi) G(\varphi', \psi') \frac{\sqrt{\sigma(\varphi, \psi) \sigma(\varphi', \psi')}}{R' R} \exp\left\{-j \frac{\omega_0}{c} (\Delta R + R' - R)\right\} d\psi d\varphi}{\iint G(\varphi, \psi)^2 \sigma(\varphi, \psi) R^{-2} d\psi d\varphi}. \quad (8)$$

АКФ, полученные по результатам моделирования и рассчитанные по выражению (8), представлены на рисунке 4.



**Рис. 4** Усредненные АКФ импульсных срезов:

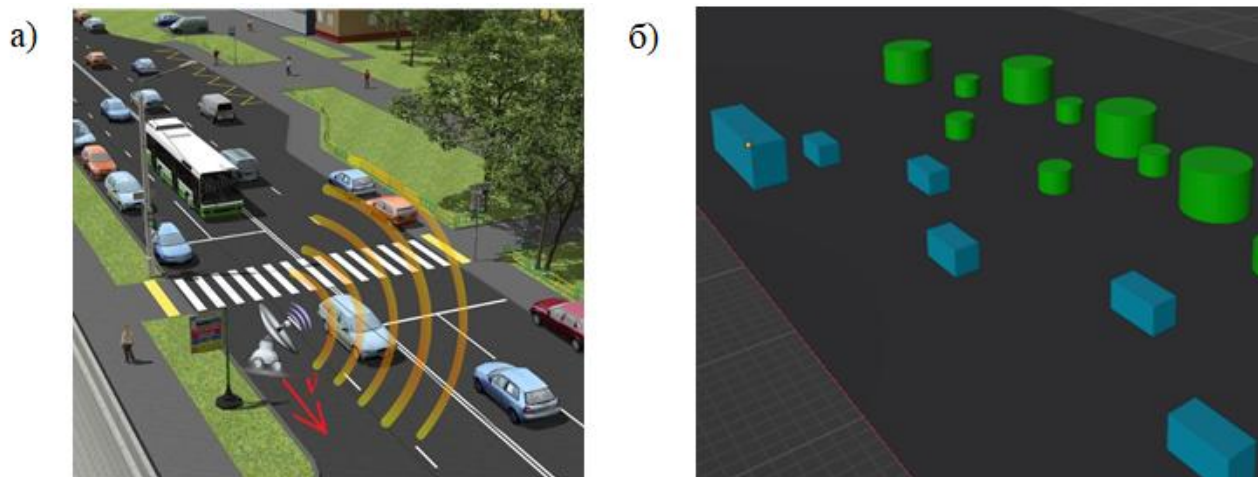
- 1 – водная поверхность (волнение 1-2 балла), 2 – эталонная АКФ (водная поверхность),  
3 – заснеженный лес, 4 – эталонная АКФ (заснеженный лес)

Полученные АКФ имеют хорошее соответствие с теоретически ожидаемыми.

Проведенные исследования разработанной модели показали соответствие результатов моделирования с результатами, полученными аналитическим путем. Представляемая модель может применяться для моделирования отраженных сигналов от сложных радиолокационных сцен с развитым рельефом подстилающей поверхности с неоднородными покровами (лес, каменистая почва, луга и т.д.). Важнейшим достоинством данной модели является учет пространственной корреляции отраженных сигналов, что дает возможность ее применения для моделирования отраженных сигналов в случае пространственно-разнесенных приемных и (или) передающих антенн. Данное свойство является особенно важным при обработке радиолокаторов, использующих технологию ММО.

Для оценки возможности моделирования сложной фоноцелевой обстановки был проведен эксперимент, целью которого являлось получение радиолокационных портретов движущихся автомобилей на фоне проезжей части (рисунок 5 а)) с помощью ЛЧМ радиолокатора, работающего в сантиметровом диапазоне длин волн.

Было выполнено моделирование радиолокационной сцены (рисунок 5 б)) и выполнена обработка полученных сигналов по алгоритму, используемому в ЛЧМ радиолокаторе.

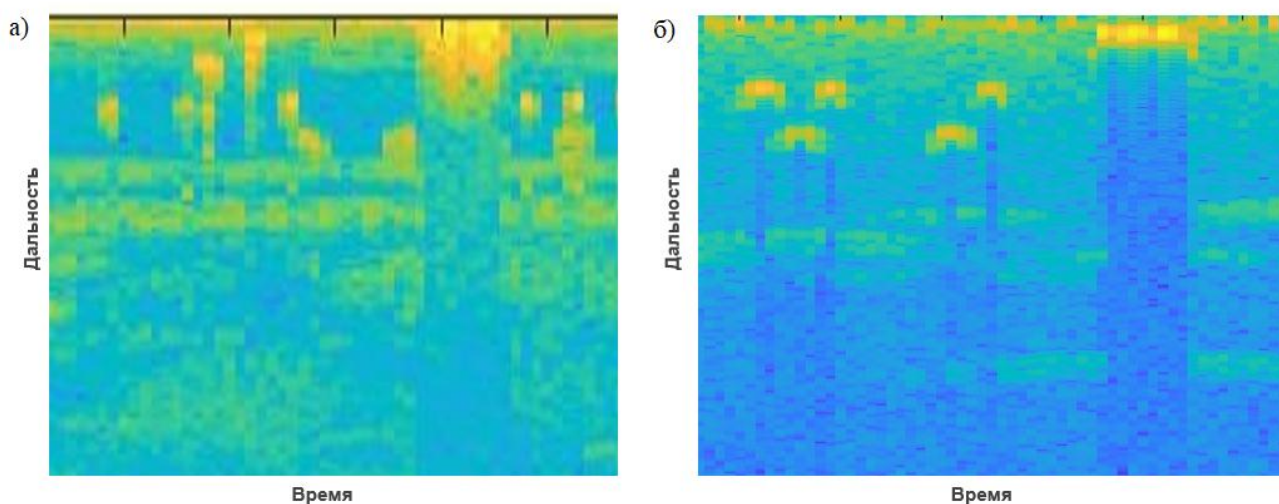


**Рис.5 Схематичное изображение эксперимента:**

а) сцена эксперимента,

б) модель сцены

Дально-временные портреты, полученные в результате эксперимента, представлены на рисунке б а), в результате моделирования – на рисунке б б).



**Рис. 6 Сравнение результатов**

а) результат эксперимента,

б) результат моделирования

Полученные результаты позволяют качественно подтвердить корректность моделирования реального сигнала, а также корректность методов, использованных при создании модели отраженных сигналов.

Полученные результаты согласуются с известными теоретическими результатами [1,3] и результатами экспериментальных исследований, проводимых в НИИИС.

### Библиография

1. Полунатурное моделирование бортовых радиолокационных систем, работающих по земной поверхности: учебное пособие / **В.Г. Важенин** [и др.]; под общ. ред. В. Г. Важенина. — Екатеринбург: Изд-во Урал. ун-та, 2015.— 208 с.
2. **Канащенков, А. И.** Формирование облика авиационных систем управления вооружением / А.И. Канащенков. — М.: Радиотехника, 2006.— 336 с.
3. **Зубкович, С. Г.** Статистические характеристики радиосигналов, отраженных от земной поверхности / С. Г. Зубкович. - М.: Сов. радио, 1968. — 224 с.

4. **Жуковский, А.П.** Теоретические основы радиовысотометрии /А.П. Жуковский, Е.И. Оноприенко, В.И. Чижов. – М.: Сов. радио, 1979. – 320 с.
5. **Фельдман, Ю.И., Мандуровский, И.А.** Теория флуктуаций локационных сигналов, отраженных распределенными целями / Фельдман Ю.И. – М.: Радио и связь, 1988. – 272 с.
6. **Анфимова, А.В., Куликов, Ю.М., Львутин, А.В.,** Математическое моделирование радиолокационного сигнала, отраженного от взволнованной морской поверхности //VII Всероссийская научно-практическая конференция: сборник трудов «Технологии разработки и отладки сложных технических систем», МГТУ им. Н.Э. Баумана, г.Москва, 2020, 454 с.
7. **Дубровин, М.С., Куликов, Ю.М., Анфимова, А.В.,** Моделирование радиолокационного сигнала, отражённого от подстилающей поверхности //Сборник материалов XXVII Международной научно-технической конференции «Информационные системы и технологии». ИСТ – 2021, НГТУ им. Р.Е. Алексеева, г. Нижний Новгород, 2021.
8. **Быков, В.В.** Цифровое моделирование в статистической радиотехнике/ Сов. радио, 1971. – 328 с
9. **Боркус, М.К.** Корреляционные измерители путевой скорости и угла сноса летательных аппаратов / М.К. Боркус, А.Е. Черный. — М.: Сов. радио, 1973.— 169 с.
10. **Кондратенков, Г.С.** Радиовидение. Радиолокационные системы дистанционного зондирования Земли: учеб. пособие для вузов / Г.С. Кондратенков, А.Ю. Фролов; под ред. Г.С. Кондратенкова. — М.: Радиотехника, 2005. — 368 с.

## СТЕНД «МИКРОСКОП» ДЛЯ ПОВЕРКИ МИКРОВОЛНОВЫХ РАДИОИНТЕРФЕРОМЕТРОВ

**Е.В. Ботов**

ORCID: 0000-0003-0263-8182 e-mail: bev.house@mail.ru

**В.Н. Иконников**

ORCID: 0000-0001-9455-4001 e-mail: ikonnikovvn@gmail.com

**Н.С. Корнев**

ORCID: 0000-0002-7848-3917 e-mail: korneff15@rambler.ru

**Е.С. Митин**

ORCID: 0000-0001-8747-2243 e-mail: evgeni-miti@yandex.ru

**А.В. Назаров**

ORCID: 0000-0002-0261-1724 e-mail: nazarov52@mail.ru

**А.А. Седов**

ORCID: 0000-0002-2489-7516 e-mail: aasedov@gmail.com

**Д.А. Трегубенко**

ORCID: 0000-0002-2067-9554 e-mail: tregubenko-dima@mail.ru

Федеральное государственное унитарное предприятие «Российский федеральный ядерный центр –  
Всероссийский научно-исследовательский институт экспериментальной физики»

Микроволновые радиоинтерферометры широко используются при исследованиях газодинамических процессов для измерения перемещений фронтов детонации и ударных волн. Технические характеристики радиоинтерферометров обеспечивают погрешность измерений не хуже  $\pm 0,15$  мм. В статье описаны устройство и принцип действия стенда «Микроскоп» для определения погрешности измерения перемещения отражающей поверхности. В качестве эталона сравнения в составе стенда используется стеклянная штриховая мера 2 класса точности. Приведена методика поверки радиоинтерферометров при помощи данного стенда, а также результаты определения погрешности радиоинтерферометра по представленной методике. Дана теоретическая оценка погрешности измерения перемещений с помощью стенда «Микроскоп».

**Ключевые слова:** микроволновый радиоинтерферометр, интерферограмма, измерение перемещений, погрешность, поверка, измерительная техника.

## «MICROSCOPE» STAND FOR VERIFICATION OF MICROWAVE RADIO INTERFEROMETERS

**E.V. Botov, V.N. Ikonnikov, N.S. Kornev, E.S. Mitin, A.V. Nazarov, A.A. Sedov, D.A. Tregubenko**

Federal state unitary enterprise «Russian federal nuclear center - All-russian scientific-research institute  
of experimental physics»

**Purpose:** The article describes the device and principle of operation of the «Microscope» stand for determining and confirming the measurement error of the reflecting surface movement.

**Design/methodology/approach:** The stand was created on the basis of an instrumental microscope BMI-1, on the slide table of which a glass line measure of the 2 accuracy class is fixed. The photodetector records changes in the luminosity of the laser, according to the registered signal of which the movement of a slide table with a reflector is determined.

**Findings:** The technique of verification of microwave radio interferometers using the developed stand is presented, as well as the results of determining the error of the radio interferometer that does not exceed  $\pm 0.15$  mm. A theoretical assessment of the measurement error of the reflecting surface movement using the stand does not exceed  $\pm 3$   $\mu$ m.

**Research limitations/implications:** The stand allows to determine the measurement errors of the movements of the reflecting surface from 0 to 100 mm.

**Originality/value:** The developed stand, unlike existing length measuring instruments, has the ability to automatically reproduce the movement of a fixed value, synchronization of registering measurement results, as well as a lower cost.

**Keywords:** microwave radio interferometer, interferogram, movement measuring, measurement error, verification, measuring equipment.



## Введение

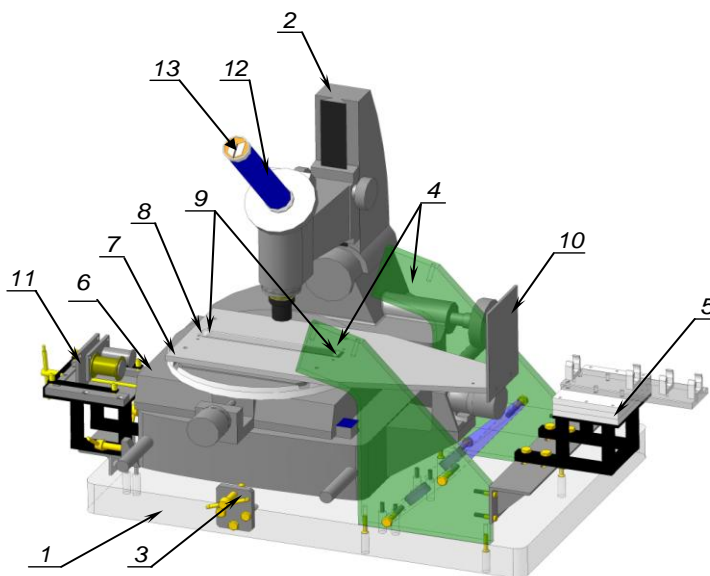
В настоящее время микроволновые радиоинтерферометры активно внедряются как в отечественную [1], так и зарубежную [2, 3] практику экспериментальных исследований быстропротекающих процессов. Использование микроволновой радиоинтерферометрии в экспериментах по исследованию процессов детонации, горения, метания позволяет получать информацию о движении объектов и газодинамических фронтов, а также динамике конструкций в реальном времени с высоким пространственным и временным разрешением, не оказывая при этом влияния на исследуемый процесс.

При решении большинства практических задач с помощью микроволновых радиоинтерферометров измеряются зависимости перемещения объектов от времени в пределах 100 мм, причем технические характеристики прибора обеспечивают погрешность измерений не хуже  $\pm 0,15$  мм. Для подтверждения указанной погрешности необходимо устройство, воспроизводящее перемещения как минимум в 3–5 раз точнее.

Имеющийся опыт применения горизонтальных длиномеров зарубежного производства [4, 5] для испытаний микроволновых радиоинтерферометров, несмотря на высокую точность, выявил ряд существенных недостатков. Помимо высокой стоимости данного оборудования, это невозможность автоматического воспроизведения перемещения фиксированной величины и отсутствие синхронизации при регистрации результатов измерений. В РФЯЦ-ВНИИЭФ разработан стенд «Микроскоп», предназначенный для определения погрешности измерения перемещения отражающей поверхности с погрешностью не более 3 мкм.

## Устройство и принцип действия стенда

Стенд «Микроскоп» создан на основе инструментального микроскопа БМИ-1. В основе измерительной части стенда лежит стеклянная штриховая мера 2 класса точности [4]. Общий вид стенда показан на рис. 1.



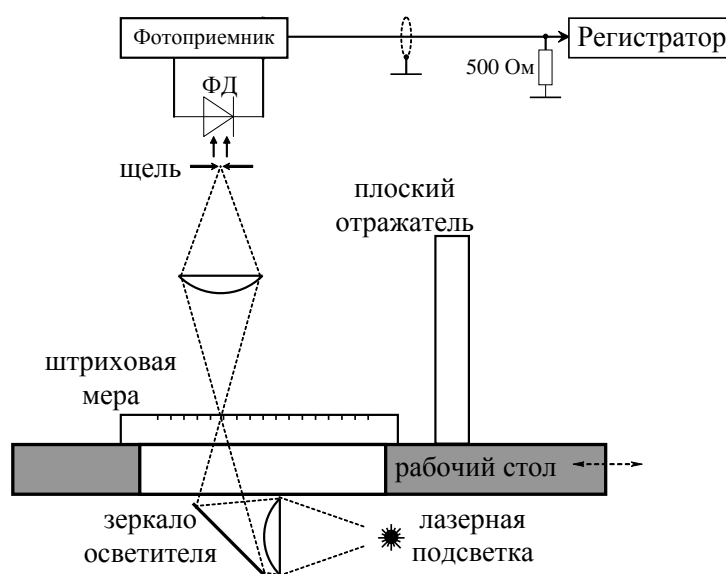
**Рис. 1. Общий вид стенда «Микроскоп»: 1 – основание; 2 – инструментальный микроскоп БМИ-1; 3 – зажимы; 4 – упоры; 5 – узел крепления облучателя; 6 – подвижный рабочий стол; 7 – опорная плита; 8 – штриховая мера; 9 – накладки; 10 – отражатель; 11 – электромагнитный узел; 12 – коллиматор; 13 – фотоприемник.**

На массивном основании, выполненном (для снижения вибрации) из толстого листа органического стекла, крепится микроскоп БМИ-1. Подвижный рабочий стол стенда имеет возможность смещаться на 150 мм. На рабочем столе закрепляется опорная плита, на которой жестко зафиксированы образцовая штриховая мера и отражатель. На упорах устанавливается узел крепления антенно-фидерной системы, на котором при испытаниях устанавливается облучатель

диэлектрического зондирующего устройства радиоинтерферометра. С боковой стороны стенда закрепляется электромагнитный узел, удерживающий стол в начальном положении.

Штатная подсветка микроскопа БМИ-1 заменена на лазерную подсветку, которая находится на неподвижной части стенда непосредственно под штриховой мерой. Расстояние между соседними штрихами меры составляет  $(1,0000 \pm 0,0025)$  мм.

При отключении электромагнита происходит движение рабочего стола по направлению к облучателю. Изображение штрихов меры при этом проецируется через оптическую систему микроскопа и коллиматор на щель. Штрихи в процессе перемещения стола периодически затевают щель. Изменения освещенности считываются при помощи фотоприемника, который размещается непосредственно над щелью. Зарегистрированная последовательность импульсов в дальнейшем используется для определения перемещения рабочей части стенда. Принцип действия стенда проиллюстрирован на рис. 2.



**Рис. 2. Принцип работы стенда «Микроскоп»**

Питание лазерной подсветки и управление электромагнитом осуществляется при помощи блока электроники, смонтированного в металлическом корпусе и соединенного с соответствующими элементами стенда кабельными жгутами. В блоке электроники предусмотрена возможность подачи и отключения питания на электромагнит от внешнего источника путем подачи сигнала с генератора синхроимпульсов.

Инструментальная погрешность измерений перемещения с помощью стенда ( $\Delta_c$ ) складывается из погрешности штриховой меры, которая в соответствии с документацией составляет 2,5 мкм, и погрешности, вносимой фотоприемником и регистрирующим осциллографом.

Погрешность, вносимая фотоприемником и осциллографом может быть вычислена по формуле:

$$\Delta X = V \cdot T_s,$$

где  $V$  – максимальная скорость движения рабочего стола стенда,  $T_s$  – суммарная временная задержка.

Погрешность, вносимая фотоприемником, связана с временной задержкой преобразования изменения интенсивности падающего на фотодиод светового потока в электрический сигнал, которая не превышает 0,135 мкс.

Погрешность, вносимая осциллографом связана с временной асинхронностью записи в память времени перекрытия светового потока штрихами меры относительно времени записи данных радиоинтерферометра и составляет не более 0,002 мкс.

Суммарная постоянная времени, таким образом, составляет не более 0,137 мкс, что при скорости движения стола станда не более 20 мм/с, добавляет дополнительную погрешность не более 0,00274 мкм. Таким образом, можно утверждать, что абсолютная инструментальная погрешность измерений перемещения отражающей поверхности с помощью станда  $\Delta_C$  не превышает по модулю 3 мкм.

### Проверка микроволновых радиоинтерферометров

Схема испытаний по оценке погрешности измерения перемещения микроволновым радиоинтерферометром показана на рис. 3. Радиоинтерферометр генерирует на выходе непрерывный зондирующий сигнал миллиметрового диапазона длин волн, который через боковой канал двухканальной волноведущей системы с направленным ответвителем поступает на облучатель. Приём отражённого сигнала обеспечивается через прямой канал волноведущей системы.

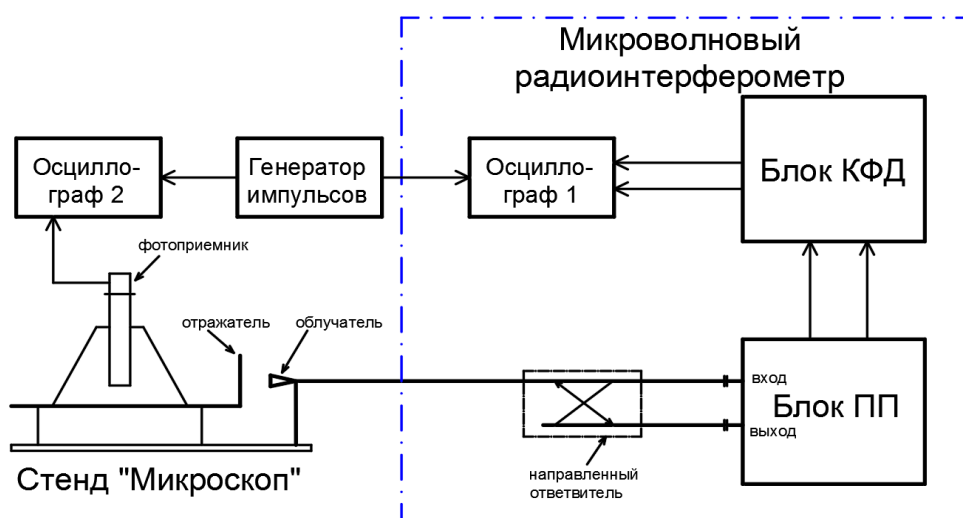


Рис. 3. Схема проверки микроволновых радиоинтерферометров

Измерение перемещений при помощи радиоинтерферометра может проводиться путем цифровой обработки как сигналов с выхода блока квадратурных фазовых детекторов (КФД), так и сигналов на промежуточной частоте (ПЧ), зарегистрированных непосредственно на выходе блока приемопередатчика (ПП).

При измерениях с использованием блока КФД интерферограмма, несущая информацию о перемещении отражателя, регистрируется непосредственно на осциллограф 1, после чего сохраняется в файл и обрабатывается специальной программой на персональном компьютере. Результатом обработки интерферограммы является X-t диаграмма, описывающая линейное перемещение отражателя в зависимости от времени.

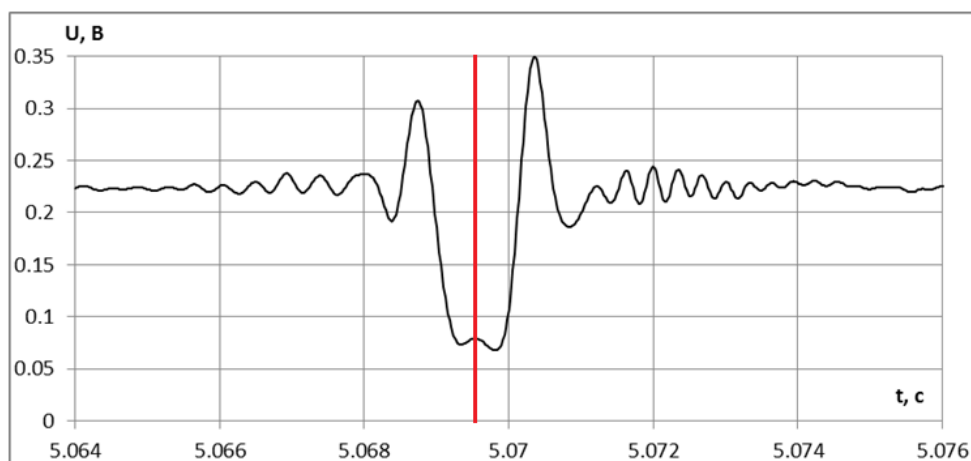
В качестве регистраторов в схеме на рис. 3 используются высокоскоростные осциллографы. Генератор импульсов (например, Г5-82) используется лишь для синхронного первичного запуска обоих осциллографов. Поскольку внутренние часы современных осциллографов имеют погрешность времени не хуже  $10^{-7}$ , синхронный запуск гарантирует синхронность времени на обоих регистраторах.

При регистрации сигналов на ПЧ осциллограф 1 необходимо использовать на максимальной частоте дискретизации (10 Гвыб/с) ввиду большого значения ПЧ (1,2 ГГц). Длительность непрерывной записи в таком режиме составляет всего несколько миллисекунд, а длительность полного прохода рабочего стола станда составляет около 12 секунд. Поэтому запись сигнала ПЧ осуществляется в режиме сегментированной регистрации: сигнал записывается небольшими отрезками длительностью несколько (от 2 до 10) наносекунд с интервалами времени между ними 3 миллисекунды. Длительность одного сегмента позволяет определить сдвиг фазы между опорным

и информационными сигналами ПЧ и провести квадратурное фазовое детектирование в цифровом виде. При этом с учетом небольшой скорости движения рабочего стола стенда, изменение этого сдвига фаз между соседними сегментами позволяет восстановить по ним интерферограмму. При таком способе регистрации один период интерферограммы состоит из 30-35 сегментов, каждый из которых усредняется до одного отсчета.

Частота следования сегментов задается генератором импульсов. Параллельно его сигнал регистрируется на осциллографе 2 синхронно с сигналом фотоприёмника, считывающего штрихи эталонной меры. При обработке сегментированного сигнала промежуточной частоты с осциллографа 1 используется сигнал генератора, зарегистрированный на осциллографе 2, чтобы интерферограмма и сигнал фотоприёмника имели общую временную привязку, соответствующую временной шкале осциллографа 2.

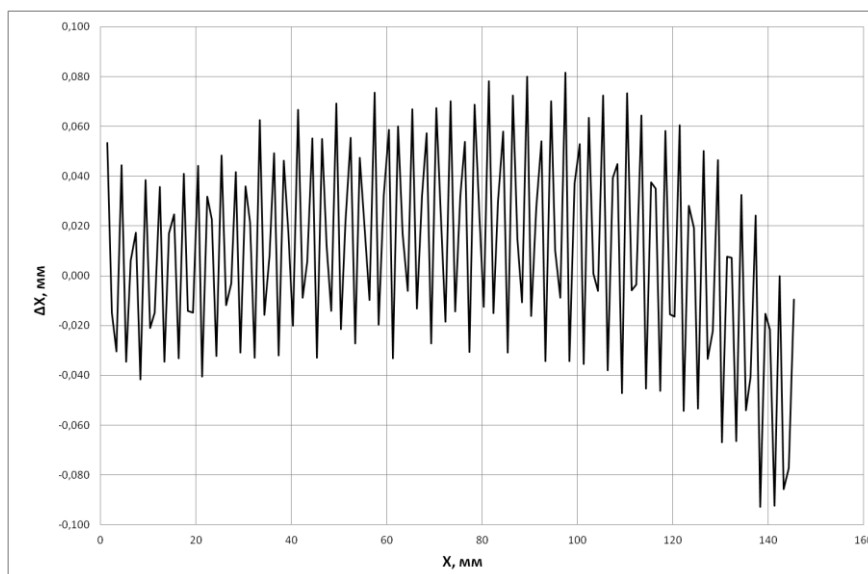
Определение перемещения по эталону осуществляется по результатам обработки сигнала, зарегистрированного на выходе фотоприёмника, фрагмент которой показан на рис. 4. Из-за дифракции монохромного излучения лазерной подсветки в окрестности каждой риски наблюдаются осцилляции интенсивности света, проходящего через оптическую щель, а в центре каждого импульса наблюдается локальный максимум, по которому и определяется момент прохождения риски.



**Рис. 4. Фрагмент осциллограммы сигнала на выходе фотоприёмника. Красная линия – момент прохождения центра риски штриховой меры.**

Для определения погрешности измерения перемещения радиоинтерферометром из  $X$ - $t$  диаграммы, полученной по данным радиоинтерферометра, делается выборка в моменты времени прохождения изображения рисков эталонной штриховой меры по оптической щели. Значение перемещения и погрешности на момент времени прохождения первой риски принимаются равными нулю, то есть первая риска принимается за начало измерения. Для последующих рисков определяется разность перемещений, полученных по данным радиоинтерферометра и сигналу фотоприёмника.

На рис. 5 приведен график зависимости погрешности измерений от перемещения отражателя, полученный по результатам обработки сигналов радиоинтерферометра с выхода блока КФД. Как видно из рис. 5, значения абсолютной погрешности не превышают по модулю 0,1 мм.



**Рис. 5. Погрешность измерения перемещения отражающей поверхности с помощью радиointерферометра**

### **Заключение**

Инструментальная погрешность стенда «Микроскоп», как следует из теоретической оценки, не превышает по модулю 3 мкм, что существенно меньше целевого значения предела допускаемой основной погрешности измерений для микроволновых радиointерферометров ( $\pm 0,15$  мм). Разработанный стенд может применяться для определения погрешности измерения перемещения отражающей поверхности микроволновыми радиointерферометрами.

### **Библиографический список**

1. Невозмущающие методы диагностики быстротекающих процессов; под ред. **А.Л. Михайлова**. – Саров: ФГУП «РФЯЦ-ВНИИЭФ», 2015. – 322 с.
2. **Mays, R.O., Tringe, J.W., Souers, P.C., Lauderbach, L.M., Baluyot, E.V., Converse, M.C., Kane, R.J.** Experimental and computational investigation of microwave interferometry for detonation front characterization. AIP Conf. Proc. 1979, 160016 (2018). doi: 10.1063/1.5045015
3. **Zhai, Z.-H., Sun, C.-L., Liu, Q., Guo, L.-W., Liu, Y.-S., Zhang, X., Wang, D.-T., Meng, K., Li, J., Du, L.-H., Liu, S.-X., Zhong, S.-C., Zhu, L.-G.** Design of terahertz-wave Doppler interferometric velocimetry for detonation physics. Appl. Phys. Lett. 116, 161102 (2020); doi: 10.1063/1.5142415
4. Длиномеры горизонтальные моделей Mini-Horizontal TELS, Horizon, Horizon Granite, Horizon Premium, THV, Labconcept/Labconcept Premium, Twinner, Alesta/Alesta Premium. Описание типа средства измерений. Приложение к свидетельству об утверждении типа средства измерений №30710
5. Приборы универсальные для измерений длины Precimar ULM 300/600/1000/1500. Описание типа средства измерений. Приложение к свидетельству об утверждении типа средства измерений №31586
6. ГОСТ 12069-90 «Меры длины штриховые брусковые. Технические условия»

## МОДЕЛИРОВАНИЕ И МАКЕТИРОВАНИЕ ОСНОВНЫХ УЗЛОВ КОНТРОЛЬНО-ПРОВЕРОЧНОЙ АППАРАТУРЫ РАДИОЛОКАЦИОННЫХ ДАТЧИКОВ

И.А. Илларионов

ORCID: 0000-0002-1065-4724 e-mail: illarionovi@list.ru

С.В. Марьевский

ORCID: 0000-0002-8763-1336 e-mail: 777serzhant888@mail.ru

Е.А. Михалицын

ORCID: 0000-0003-3144-3704 e-mail: mihalitsynea@gmail.com

А.В. Назаров

ORCID: 0000-0002-0261-1724 e-mail: aNazarov@niis.nnov.ru

Р.Р. Османов

ORCID: 0000-0002-4324-7623 e-mail: osmanov22ruslan@mail.ru

И.В. Осовицкая

ORCID: 0000-0003-3522-0225 e-mail: oivnn121@mail.ru

Б.Ю. Царев

ORCID: 0000-0001-8975-1588 e-mail: b.yu.tsarev@gmail.com

Федеральное государственное унитарное предприятие «Российский федеральный ядерный центр – Всероссийский научно-исследовательский институт экспериментальной физики»

При разработке и производстве радиолокационных датчиков (РД) возникает необходимость контроля их основных характеристик. При этом переход 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 является в настоящее время весьма актуальной. В статье представлены результаты моделирования и макетирования основных узлов КПА РД – безэховой камеры, линии задержки и узла ослабления (аттенюатора), которые в дальнейшем позволят сформировать окончательный технический облик аппаратуры КПА с учетом технических требований, предъявляемых к РД.

**Ключевые слова:** радиолокационный датчик, контрольно-проверочная аппаратура, безэховая камера, линия задержки, узел ослабления, моделирование.

## SIMULATION AND PROTOTYPING OF CONTROL-TESTING EQUIPMENT GENERAL UNITS FOR RADAR DETECTORS

I.A. Illarionov, S.V. Maryevskiy, E.A. Mihalitsyn, A.V. Nazarov, R.R. Osmanov, I.V. Osovitskaya, B.Yu. Tsarev

Federal state unitary enterprise «Russian federal nuclear center - All-russian scientific-research institute of experimental physics»

**Purpose:** This article presents the results of simulation and prototyping of control-testing equipment (CTE) general units for centimeter wavelength range radar detectors (RD) for the purpose to select the most acceptable technical solution.

**Design / methodology / approach:** A pass-through delay line model was proposed as a CTE working alternative. The results of anechoic chamber (AC) simulation, designed to simulate the propagation of the RD probe signal, the delay line prototyping, which allows controlling the RD range, and attenuation node prototyping, which serves to assess the RD energy potential, were analyzed.

**Findings:** The CTE technical appearance and structure have been determined. The AC geometry was calculated. A set of coaxial cables for the delay line implementation of the test system was determined and the base length estimated calculation was carried out taking into account the RD signal propagation delay in free space. Attenuators were selected in accordance with the required system parameters.

**Research limitations/implications:** In the future, it is possible to modernize the considered CTE to build a more complex switchable test system with simulated parameters.

**Originality/value:** The considered CTE model can be used to test the promising RD working capacity.

**Keywords:** radar detector, control-testing equipment, anechoic chamber, delay line, attenuation unit, simulation.

При разработке и производстве РД возникает необходимость контролировать их основные характеристики в лабораторных условиях. При этом переход разработок в серийное производство требует оптимизации и автоматизации процедуры проверки параметров РД. Таким образом, задача разработки КПА для первичного контроля характеристик больших партий РД является в настоящее время весьма актуальной [1].

Целью работы является разработка технического решения для системы тестирования РД, работающего в сантиметровом диапазоне длин волн.

Стандартная измерительная аппаратура не позволяет в полном объеме контролировать основные параметры таких устройств (потенциал системы – дальность действия и точность измерения дальности), что заставляет разработчиков конструировать и изготавливать сложные комплексы КПА для имитации задержки и ослабления сигнала РД [2], [3].

Макетирование и моделирование таких узлов КПА как БЭК, ЛЗ и узел ослабления (аттенуатор) позволит сформировать окончательный технический облик аппаратуры КПА и состав входящих в нее средств с учетом заданных технических требований.

### Модель КПА

КПА осуществляет канализацию высокочастотного сигнала РД, прием сигнала, обработку спектра и переизлучение сигнала в направлении РД. Условно можно выделить следующие типы систем тестирования (СТ) РД, различающиеся по радиотехническим методам построения и принципам функционирования:

- 1) СТ с фиксированными параметрами;
- 2) СТ с переключаемыми параметрами и доплеровским смещением;
- 3) СТ с имитацией параметров;
- 4) аналоговые радиопотонные СТ;
- 5) цифровые СТ.

Представленная классификация систем тестирования является условной, поскольку в коммерческих продуктах может применяться симбиоз из отдельных модулей и блоков из разных типов систем.

Для КПА РД сантиметрового диапазона длин волн предлагается использовать модель с фиксированной задержкой и ослаблением (рисунок 1) с перспективой ее дальнейшей модификации для построения более сложных переключаемых тестовых систем.

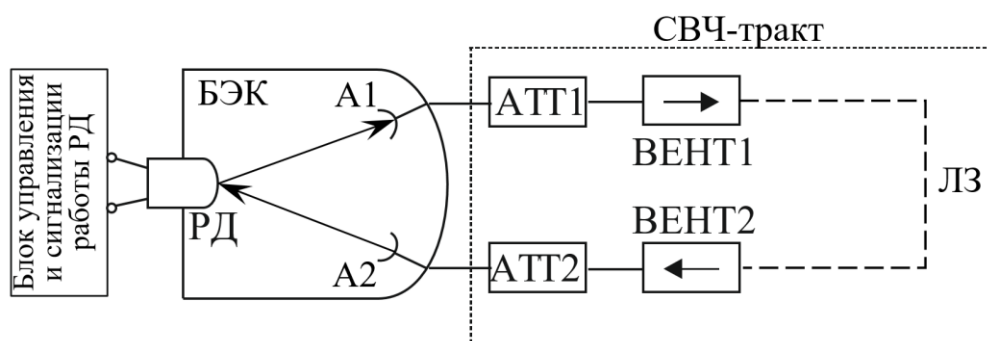


Рис. 1 Блок-схема проходной модели КПА

Рассматриваемая модель системы тестирования состоит из БЭК с установленным РД, приемной и передающей зондовых антенн (А1, А2), нагруженных на СВЧ-тракт. СВЧ-тракт состоит из аттенуаторов (АТТ1, АТТ2), вентилях (ВЕНТ1, ВЕНТ2), ЛЗ.

Работа стенда КПА происходит следующим образом. Сигнал СВЧ, излучаемый макетом РД, проходя БЭК, через приемную антенну А1 поступает в СВЧ-тракт. В приемном канале СВЧ-тракта сигнал, ослабившись в АТТ1, поступает через вентиль ВЕНТ1 в ЛЗ,

имитирующую различные дальности срабатывания РД. Далее ослабленный и задержанный относительно зондирующего сигнала поступает в передающий канал СВЧ-тракта, где, пройдя через вентиль ВЕНТ2, дополнительно ослабевает на аттенуаторе АТТ2 (комбинация АТТ1 и АТТ2 (либо один переменный аттенуатор) необходима для оценки энергетического потенциала РД). После этого сигнал с передающей антенны А2, проходя БЭК, принимается в РД. В зависимости от уровня и задержки пришедшего сигнала блок управления и сигнализации работы РД сообщает о работоспособности датчика.

### Безэховая камера

БЭК предназначена для имитации среды распространения ЛЧМ-зондирующего сигнала РД и его отражения от точечной цели.

Внешний вид предлагаемой БЭК показан на рисунке 2. БЭК имеет полуцилиндрическую форму, фанерные стенки которой изнутри обшиты радиопоглощающим материалом (РПМ). Антенная система (АС) РД располагается в центре плоскости, параллельной оси полуцилиндра. Зондовые антенны симметрично располагаются на боковой поверхности полуцилиндра в дальней зоне АС РД. Для улучшения развязки между зондовыми антеннами в БЭК устанавливается фланец – параллелепипед из соединенных между собой двух листов РПМ. Цилиндрическая форма БЭК обеспечивает равенство расстояния от зондов до АС РД при изменении углового положения зондов, что необходимо для проверки работоспособности РД в секторе рабочих углов диаграммы направленности (ДН) АС. С точки зрения исключения первого наиболее интенсивного отраженного сигнала от поверхности БЭК цилиндрическая форма не является оптимальной, так как при расположении АС РД в центре плоскости, параллельной оси полуцилиндра, всегда существует паразитный сигнал, однократно отраженный от стенки БЭК. Однако при соответствующем выборе РПМ этот сигнал можно сильно подавить, обеспечив необходимое значение КБЭ.

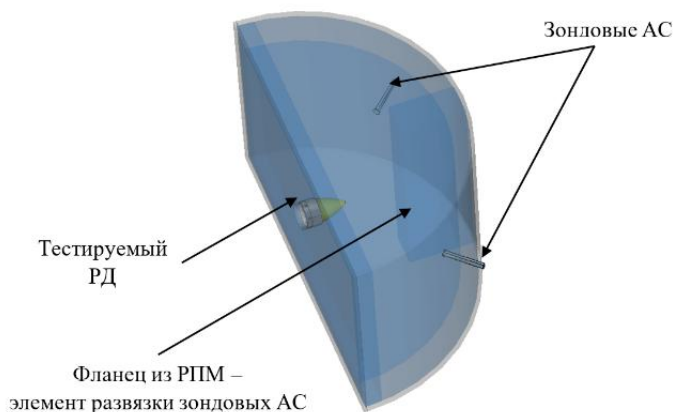
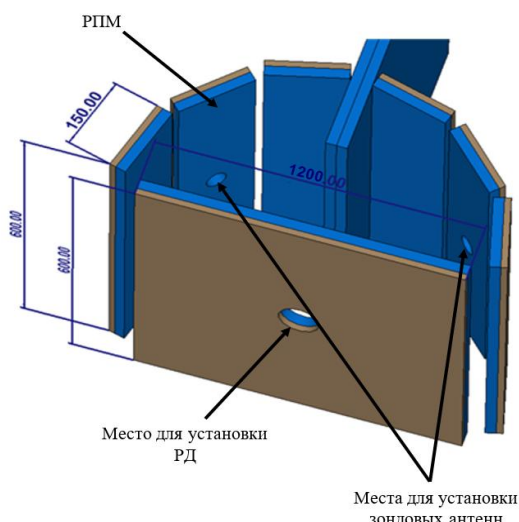


Рис. 2 Модель БЭК

Для экспериментальной отработки конструкции БЭК (выбор габаритов, типа РПМ) при известных параметрах ЛЗ, соединяющей зондовые антенны и предназначенной для имитации среды распространения ЛЧМ-зондирующего сигнала и его отражения от точечной цели, предлагается создать макет БЭК (рисунок 3).

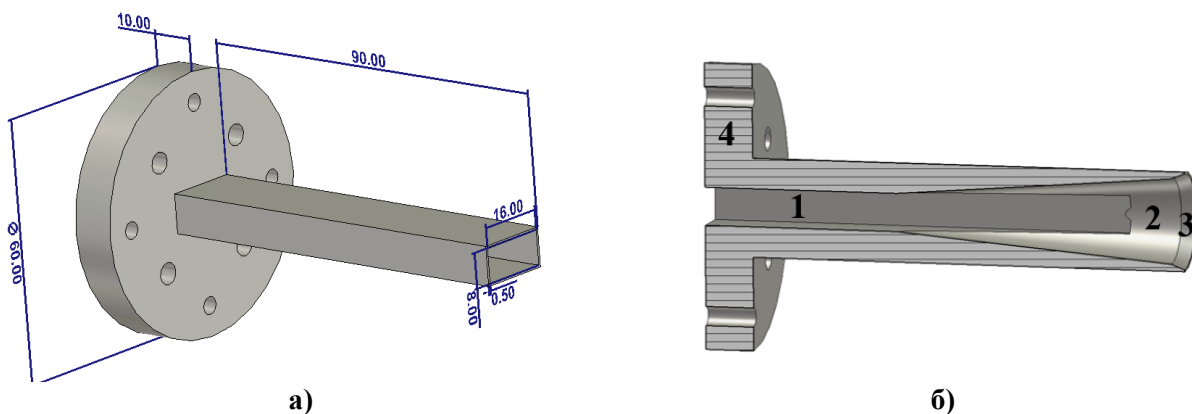




**Рис. 3 Конструкция макета БЭЖ**

Для создания КПА предлагается использовать два варианта зондовых антенн: слабонаправленные зондовые антенны на основе ОКВ прямоугольного сечения, направленные зондовые антенны на основе РПИ (рисунок 4). Выбор зондовых антенн в виде волноводных антенн обусловлен несколькими факторами: конструктивной простотой антенн, возможностью относительно точного моделирования характеристик подобных антенн, что необходимо при использовании антенн в качестве эталонных или технологических. Выходы зондовых антенн реализуются в виде коаксиальных SMA-разъемов с помощью коаксиально-волноводных переходов ПКВ1-13Р-16×8 производства фирмы «Микран» или их аналогов.

Зондовая антенна на основе ОКВ имеет по сравнению с зондовой антенной на основе РПИ меньшие ЭПР, КУ и, следовательно, меньшее апертурное взаимодействие с АС РД и меньшую площадь засветки БЭЖ. Наличие двух типов зондовых антенн позволит выбрать оптимальный вариант для заданной БЭЖ.

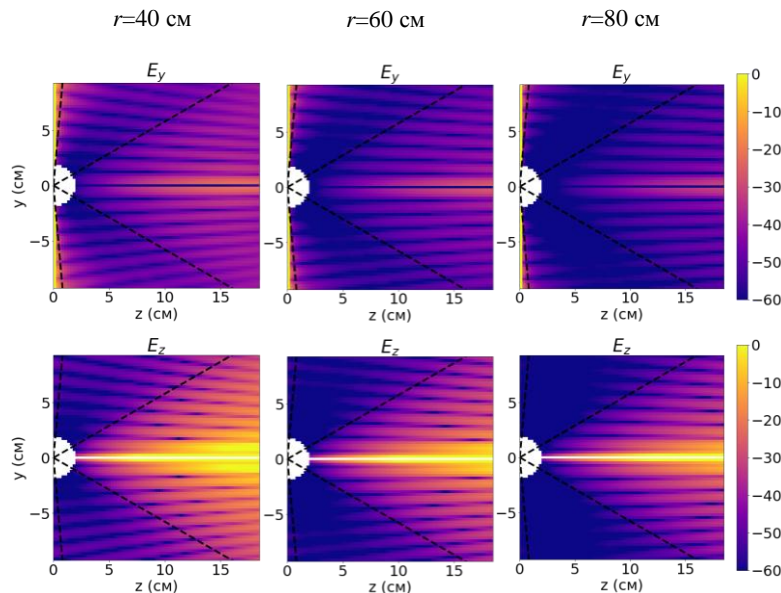


**Рис. 4 Зондовая антенна в виде открытого конца прямоугольного волновода (а); зондовая антенна в виде расфазированного рупора с изломом образующей (1 - прямоугольный волновод с сечением (16×8) мм; 2 - первая коническая переходная область; 3 - вторая коническая область; 4 - крепежный фланец) (б)**

На рисунке 5 представлены пространственные распределения КБЭ для разных поляризаций электрического поля при разных расстояниях  $r$  между зондовыми антеннами в виде РПИ и АС РД. Коэффициент безэховости (КБЭ) определялся для каждой поляризации электрического поля следующим образом:

$$\text{КБЭ (дБ)} = 20 \lg \frac{|E_1 - E_0|}{|E_0|}, \quad (1)$$

где  $E_0$  – напряженность электрического поля, создаваемая тестируемой АС РД в свободном пространстве (падающее поле), а  $E_1$  – суммарная напряженность электрического поля тестируемой АС РД и поля, рассеянного от двух симметрично расположенных зондовых антенн. Искомые напряженности в формуле (1) определялись путем численного моделирования асимптотическим методом расчета в САПР трехмерного электромагнитного моделирования. Пунктирными черными линиями на рисунках показаны направления излучения тестируемого РД с расчетным КУ 0 дБ.



**Рис. 5** Пространственные распределения КБЭ для разных поляризаций электрического поля при разных расстояниях между зондовыми АС РРИ и АС РД

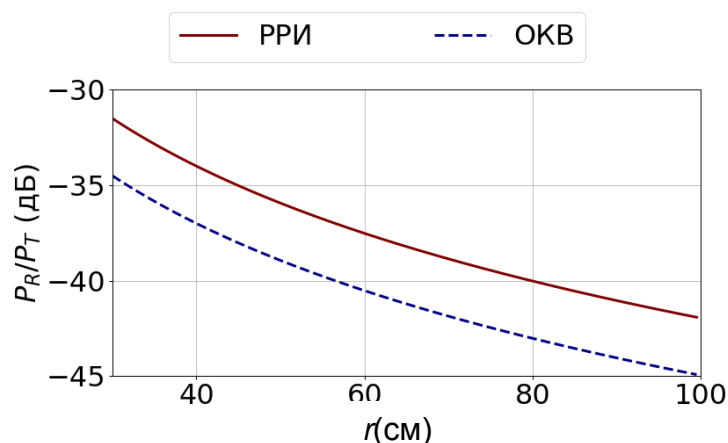
Как следует из рисунка 5, значение КБЭ при расстояниях между зондовыми антеннами и РД более 60 см менее минус 60 дБ в рабочих зонах, соответствующих рабочим секторам ДН АС РД. При расстоянии 40 см значение КБЭ – более минус 40 дБ.

Оценка влияния зондовых антенн на качество КПА РД при расстоянии 40 см между АС зондов и АС РД, проведенная в САПР трехмерного электромагнитного моделирования, показала, что развязка антенн на основе ОКВ с запасом удовлетворяет предъявляемым требованиям. Для зондов на основе РРИ требуется дополнительное введение радиопоглощающего элемента (например, фланца из соединенных между собой двух листов РПМ, как в предлагаемом макете БЭК), позволяющего улучшить развязку.

Уровень затухания при прямом прохождении сигнала от АС РД до приемной зондовой антенны можно оценить по формуле

$$\frac{P_R}{P_T} [\text{дБ}] = G_T [\text{дБ}] + G_R [\text{дБ}] + 20 \lg \frac{\lambda}{4\pi r}, \quad (2)$$

где  $P_T$  – мощность сигнала, излучаемого (принимаемого) РД,  $P_R$  – мощность сигнала, принимаемого (излучаемого) зондовой антенной,  $G_T$  – коэффициент усиления (КУ) АС РД,  $G_R$  – КУ зондовой антенны,  $\lambda$  – длина волны излучения,  $r$  – расстояние между АС РД и зондовой антенной. На рисунке 6 приведена зависимость уровня сигнала на выходе зондовой антенны по отношению к сигналу на входе РД от расстояния  $r$  между зондовой антенной и АС РД (длина волны в свободном пространстве  $\lambda = 2$  см) для разных зондовых антенн.



**Рис. 6** Уровень сигнала на входе (выходе) зондовой антенны по отношению к сигналу на выходе (входе) РД в зависимости от расстояния  $r$  между зондовой антенной и АС РД

### 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. Поэтому для исследуемого РД предлагается реализация ЛЗ сантиметрового диапазона длин волн, которая может быть выполнена на коаксиальных кабелях.

В качестве опытного варианта возможно использование кабельных сборок ф. Mini-Circuits на прецизионных 50-омных кабелях с заявленной производителем повышенной стабильностью вносимых потерь ( $|S_{21}|$ ), КСВн и ФЧХ к изгибам кабеля с радиусом до 76,2 мм.

Опытные кабельные сборки:

1. CBL-1M-SMNM+ (2 шт. – сер. № 204203, сер. № 204206): sma(m)-N(m), длина 1 м, типичные значения  $|S_{21}|=2,06$  дБ (15 ГГц), КСВн=1,1 (15 ГГц);
2. CBL-3FT-SFNM+ (2 шт. – сер. №199782, сер. № 199790): sma(f)-N(m), длина 0,91 м, типичные значения  $|S_{21}|=1,9$  дБ (15 ГГц), КСВн=1,1 (15 ГГц).

С целью подтверждения заявленных характеристик и оценки возможности применения в макете КПА произведены измерения характеристик вносимых потерь  $|S_{21}|$ , КСВн и ФЧХ указанных кабельных сборок (таблица 1).

**Таблица 1** – Измеренные значения характеристик кабельных сборок

Кабельная сборка	Потери $ S_{21} $ , дБ	КСВн	Задержка сигнала $\tau$ , нс
1 (№204203)	2	1,16	4,96
1 (№204206)	1,97	1,16	4,97
2 (№199782)	1,75	1,1	4,56
2 (№199790)	1,91	1,12	4,45

Результаты, полученные в ходе экспериментальных исследований, позволяют сделать вывод о возможности применения кабельных сборок из вышеуказанного типа кабеля в качестве линии задержки в макете стенда КПА для РД сантиметрового диапазона длин волн. Для имитации задержки с учетом уставки срабатывания РД могут подходить различные комбинации сборок, для их соединения между собой дополнительно необходимо применять ВЧ-переходы.

## Узел ослабления (аттенюатор)

Для дополнительной оценки потенциала РД необходим аттенюатор (группа аттенюаторов), суммарное ослабление которого совместно с БЭК и СВЧ-трактом должно перекрывать значение потенциала, заложенное в технических требованиях к РД. В качестве опытного варианта возможно использование прецизионных 50-омных аттенюаторов ф. Mini-Circuits с заявленной производителем повышенной стабильностью вносимого ослабления  $|S_{21}|$ , КСВн.

В частности, для оценки потенциала РД сантиметрового диапазона длин волн в диапазоне уставки срабатывания необходимо использовать комбинацию аттенюаторов BW-S6W20+ ( $|S_{21}|=6$  дБ) и BW-S20W20+ ( $|S_{21}|=20$  дБ) (для ОКВ) или одиночный аттенюатор BW-S30W20+ ( $|S_{21}|=30$  дБ) (для РПИ).

С целью подтверждения заявленных характеристик и оценки возможности применения в качестве узла ослабления в стенде КПА проведены измерения параметра  $|S_{21}|$  и КСВн указанных аттенюаторов (таблица 2).

Таблица 1 – Измеренные значения характеристик кабельных сборок

Аттенюатор	BW-S6W20+	BW-S20W20+	BW-S30W20+
КСВн	1,03	1,09	1,12
Ослабление, дБ	6,11	19,82	29,74

Проведенные исследования аттенюаторов также свидетельствуют о высоком качестве их характеристик и принципиальной возможности их применения при изготовлении макета КПА.

В перспективе предлагается использовать управляемый электромеханический аттенюатор Д6М-18-11Р ф. «Микран» с максимальной рабочей частотой до 18 ГГц, диапазоном перестройки (0-81) дБ с шагом 1 дБ. Данное решение особо удобно в случае проверки серийно изготавливаемых РД, т.к. встроенная система команд позволит автоматизировать измерения с возможностью дистанционного управления с ПК через интерфейсы LAN, USB, RS-232.

При мелкосерийном производстве РД возможен вариант использования ступенчатого аттенюатора AF9003-69-32 ф. «API/Weinschel» с максимальной рабочей частотой до 18 ГГц, диапазоном перестройки (0-69) дБ с шагом 1 дБ, предельной мощностью до 2 Вт.

## Вентили

Вентили в каналах СВЧ-тракта КПА (рисунок 1) предназначены для организации проходной схемы имитации задержки сигнала РД, благодаря чему один канал до ЛЗ (А1 – ЛЗ) работает на прием, второй (ЛЗ – А2) – на передачу. Также указанные узлы помогают дополнительно исключить попадание отраженного сигнала из тракта СВЧ после вентиляй обратно в БЭК.

Предлагается использовать вентили D4I8018 ф. «DiTomMicrowave» с максимальной рабочей частотой до 18 ГГц, развязкой – 32 дБ, КСВн – 1,45.

## Выводы

Таким образом, определен технический облик и состав КПА. Рассчитана геометрия БЭК, выбраны типы зондирующих антенн. Определен комплект коаксиальных кабелей для реализации ЛЗ, выбраны вентили и аттенюаторы в соответствии с требуемыми параметрами системы. Проведенные исследования также свидетельствуют о возможности применения данных узлов в макете КПА РД.

В перспективе, в случае положительных результатов при апробации, возможна дальнейшая модернизация рассматриваемой КПА для построения более сложной переключаемой тестовой системы РД с имитацией параметров.

#### **Библиографический список**

1. **Дорогов, А.Д., Марьевский, С.В., Назаров, А.В., Османов, Р.Р.**, Технический облик стенда контрольно-проверочной аппаратуры для радиолокационных датчиков // Информационные системы и технологии. ИСТ – 2021: Материалы XXVII Международной научно-технической конференции. – Н.Новгород, 2021. – С. 993-997.
2. **Дмитриев, Н. И., Бабковский, А. П., Селезнев, Н. Е.** Контрольно-проверочная аппаратура для маловысотных радиоальтиметров. – 16-я Международная крымская конференция «СВЧ техника и коммуникационные технологии». Материалы конференции. [Севастополь 11-15 сентября 2006 г.]. Севастополь: Вебер, 2006, стр. 879-881.
3. **S. Heuel, D. McCarthy.** Real-time radar target generation // Microwave Journal. 2015.

## ОСОБЕННОСТИ КОРРЕЛЯЦИОННОГО ОБНАРУЖЕНИЯ РАСПРЕДЕЛЕННЫХ ОБЪЕКТОВ

К.Д. Пирогова

ORCID: 0000-0002-1232-2125 e-mail: [ksyuhapirogova@mail.ru](mailto:ksyuhapirogova@mail.ru)

Нижегородский государственный технический университет им. Р.Е. Алексеева

Одним из направлений развития современных радиолокационных систем является использование многопозиционной конфигурации, чем обеспечивается улучшение характеристик системы в части эффективности обнаружения малозаметных целей, повышения разрешающей способности и точности. В многопозиционных радиолокационных системах в условиях воздействия нестационарных помех целесообразно использовать корреляционный алгоритм обнаружения. Этот алгоритм основан на вычислении взаимно-корреляционного момента между процессами, действующими в пересекающихся лучах двух или более радаров и сравнение его с порогом. В статье рассматриваются особенности применения корреляционного алгоритма при обнаружении распределенных целей, которые занимают одновременно несколько элементов разрешения. Методом математического моделирования получены характеристики обнаружения. Проведен сравнительный анализ эффективности обнаружения распределенных и точечных целей на фоне гауссовского шума, а также оценен выигрыш в отношении сигнал-шум по сравнению со случаем обнаружения целей одиночным радаром.

**Ключевые слова:** радиолокационные системы обнаружения, многопозиционные РЛС, корреляционный алгоритм, распределенная цель

## FEATURES OF CORRELATION DETECTION OF DISTRIBUTED OBJECTS

K.D. Pirogova

Nizhny Novgorod state technical university n. a. R.E. Alexeev

**Purpose:** The article discusses the features of using the correlation algorithm when detecting distributed targets that simultaneously occupy several resolution elements.

**Design / methodology / approach:** The article simulates a multisite radar system consisting of two radars. The input data for the correlation procedure are the results of monostatic target detection in single sensors and the complex samples of 2D FFT after the beam forming. The correlation procedure is performed separately within each area of 2D FFT occupied by a distributed target. Further, the obtained value of the correlation moment is compared with the threshold and a decision is made about the presence or absence of the target.

**Findings:** A comparative analysis of the efficiency of detection of distributed and point targets against the background of Gaussian noise was carried out, and the gain in signal-to-noise ratio was estimated compared to the case of target detection by a single radar.

**Research limitations/implications:** The research did not take into account the change in the position of the 2D FFT maximum from the same target for each radar in a distributed radar system.

**Originality/value:** The value of this work lies in increasing the value of the probability of correct detection of a distributed target. The article also notes that the size of the area occupied by the target does not affect the detection characteristics, which in turn indicates that the detection probability is determined by the 2D FFT maximum.

**Key words:** radar system, multisite radar system, correlation algorithm, distributed target

## Введение

Современное повышение требований к основным характеристикам радиолокационных станций (дальность действия, точность измерений, эффективность защиты от помех) привело к необходимости разработки таких систем, которые должны им соответствовать. Одним из усовершенствованных принципов построения систем обнаружения целей является переход от отдельных радиолокационных станций с одной передающей и одной приемной позициями к многопозиционным радиолокационным станциям. Такие станции состоят из разнесенных в пространстве передающих и приемных позиций. Использование многопозиционной РЛС позволяет расширить зону обнаружения, обеспечить лучшее подавление помех, повысить разрешающую способность и точность измерения координат [1].

### Корреляционный алгоритм обнаружения

Использование двух и более разнесенных в пространстве радаров, которые наблюдают за одним и тем же объектом приводит к улучшению характеристик обнаружения. Использование корреляционного алгоритма в таких системах является эффективным для обнаружения слабых целей в условиях действия нестационарных помех. Этот алгоритм основан на вычислении взаимно-корреляционного момента между процессами, действующими в пересекающихся лучах двух или более радаров и сравнение его с порогом. Кросс-корреляции между двумя областями двумерного быстрого преобразования Фурье вычисляется следующим образом:

$$R(\Delta n, \Delta k) = \sum_{n=1}^N \sum_{k=1}^K \bar{S}_1(n, k) \bar{S}_2^*(n + \Delta n, k + \Delta k), \quad (1)$$

где  $\bar{S}_1(n, k)$  – матрица отсчетов в определённой области, близкой к обнаруженному локальному максимуму БПФ в одном канале;

$\Delta n, \Delta k$  – взаимные сдвиги областей двумерного БПФ, выраженные в номерах элементов разрешения результата двумерного БПФ.

При одинаковых приемных каналах:

$$\bar{S}_2^*(n, k) = \bar{S}_1^*(n, k) \exp(j\Delta\varphi), \quad (2)$$

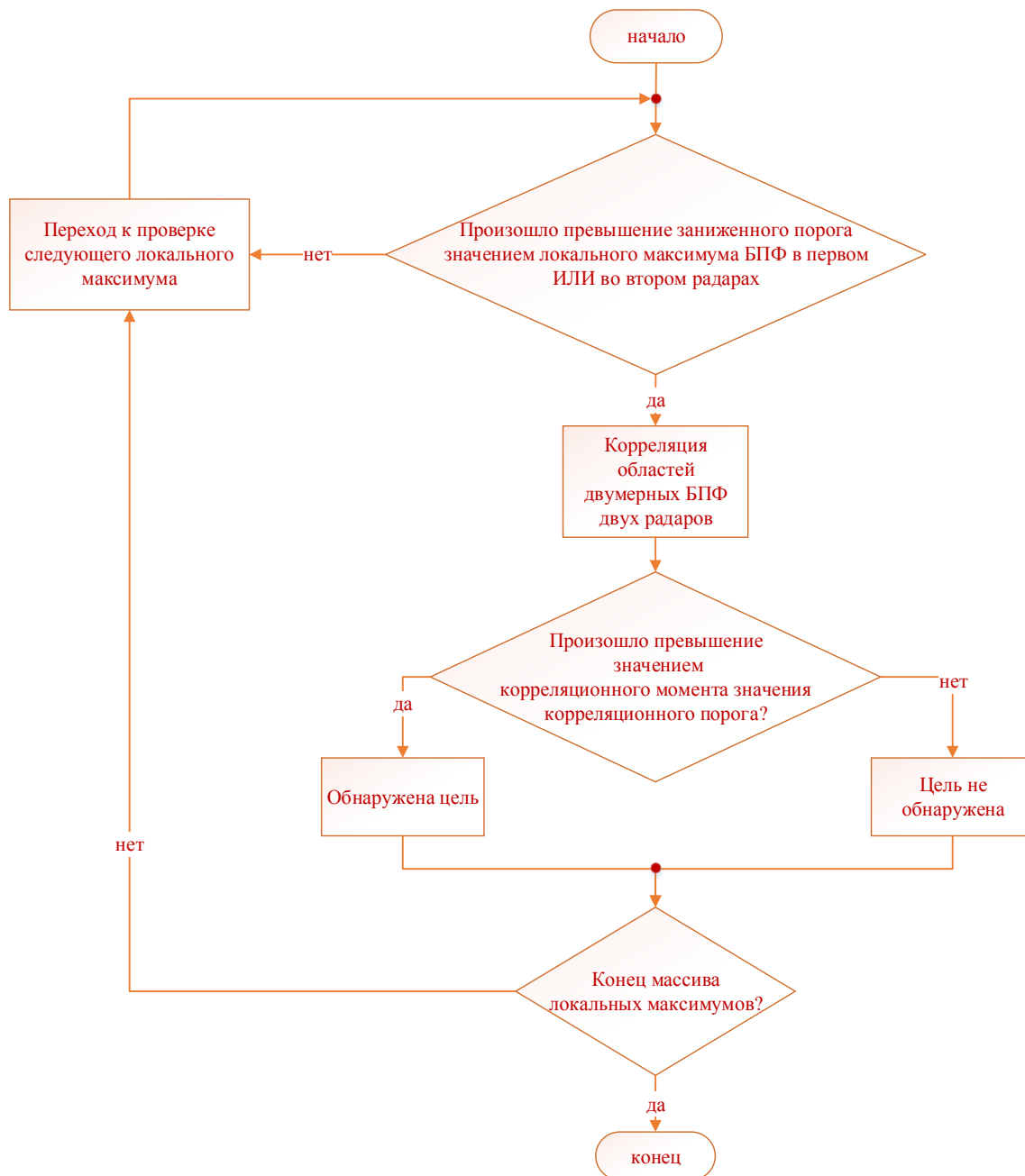
где  $\Delta\varphi$  – фазовый сдвиг между каналами.

Для случая нулевых взаимных сдвигов уравнение принимает вид:

$$R = \sum_{n=1}^N \sum_{k=1}^K |\bar{S}_1(n, k)|^2 \exp(j\Delta\varphi) = kE_s \exp(j\Delta\varphi), \quad (3)$$

где  $E_s$  – энергия сигнала.

Если два распределенных в пространстве радара наблюдают слабую цель, а требуемая вероятность ложной тревоги  $F = 10^{-6}$ , то вероятность правильного обнаружения слабой цели слишком мала. Такая ситуация может возникнуть в нашем случае, если один радар используется как эталон для дальнейшего сравнения полученных результатов. Поэтому предлагается использовать двухэтапный алгоритм обнаружения (рисунок 1). Использование двух радаров и более позволяет нам проводить предварительное обнаружение с большей вероятностью ложной тревоги  $F = 10^{-3}$ , но повышенным значением вероятности правильного обнаружения. Затем логическим блоком «ИЛИ» объединяются результаты процедуры предварительного обнаружения для дальнейшего улучшения вероятности обнаружения с последующим незначительным увеличением частоты ложных срабатываний. Процедура кросс-корреляции запускается только тогда, когда происходит обнаружение хотя бы в одном из радаров.



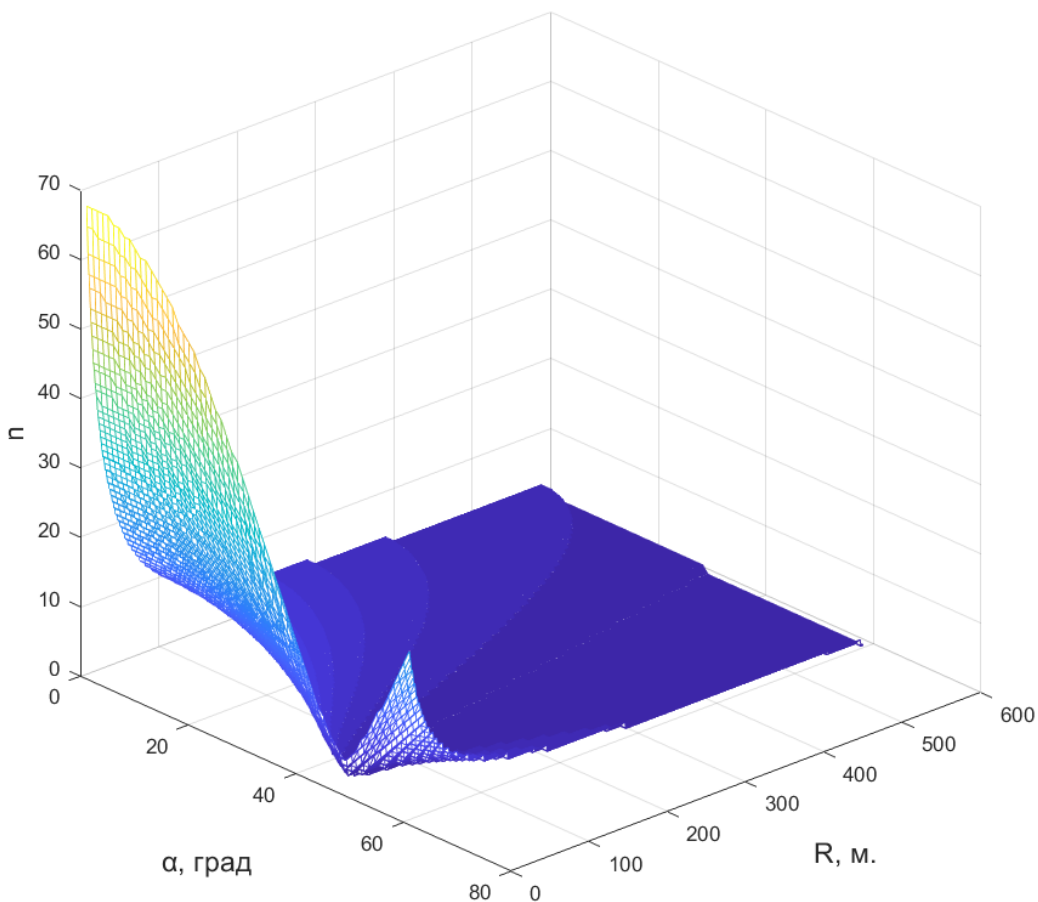
**Рис. 1. Двухэтапный алгоритм обнаружения**

### **Обнаружение распределенных объектов корреляционным алгоритмом**

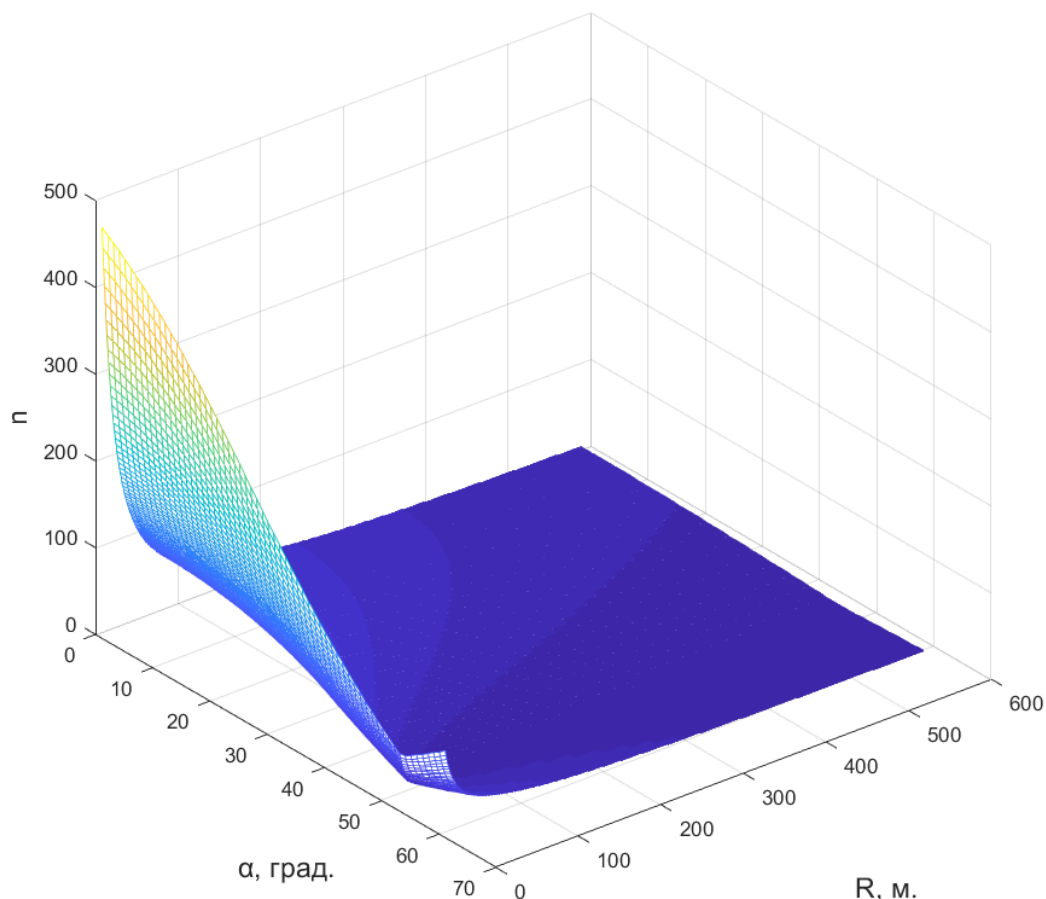
В работе [2] рассматривались преимущества корреляционного алгоритма для обнаружения точечных целей, но не оценивались характеристики обнаружения для обнаружения распределенной (многоточечной) цели. Для обнаружения распределенных целей предлагается корреляция площадок результатов БПФ двух радаров, занимаемые данными целями. Размер площадок не должен быть слишком большим, так как это приводит к повышению требуемой вычислительной мощности. Определим количество элементов разрешения результата БПФ по размерности доплеровских частот, занимаемое распределенной целью. Для этого построим графики зависимости количества занимаемых элементов разрешения результата БПФ по размерности доплеровских частот от азимутального направления и дальности до центра распределенной цели. Результаты моделирования для цели длиной 5 м и цели длиной 20 м показаны на рисунке 2 и рисунке 3 соответственно. Эффективное использование корреляционного подхода возможно в условиях так называемой зоны высокой корреляции для многоточечной цели (дальность до цели свыше 100м). Ограничиваясь этой зоной, по полученным графикам видно, что



распределенная цель может занимать от двух до пятидесяти элементов разрешения результата БПФ по размерности доплеровских частот. Руководствуясь принципом минимизации вычислительной мощности алгоритма будем в дальнейшем проводить исследования для площадок размерами 2x2 и 4x4.



**Рис. 2. Количество элементов БПФ по размерности доплеровских частот, занимаемые распределенной целью в зависимости от угла направления траектории цели и расстояния до цели (длина цели 5 м.)**



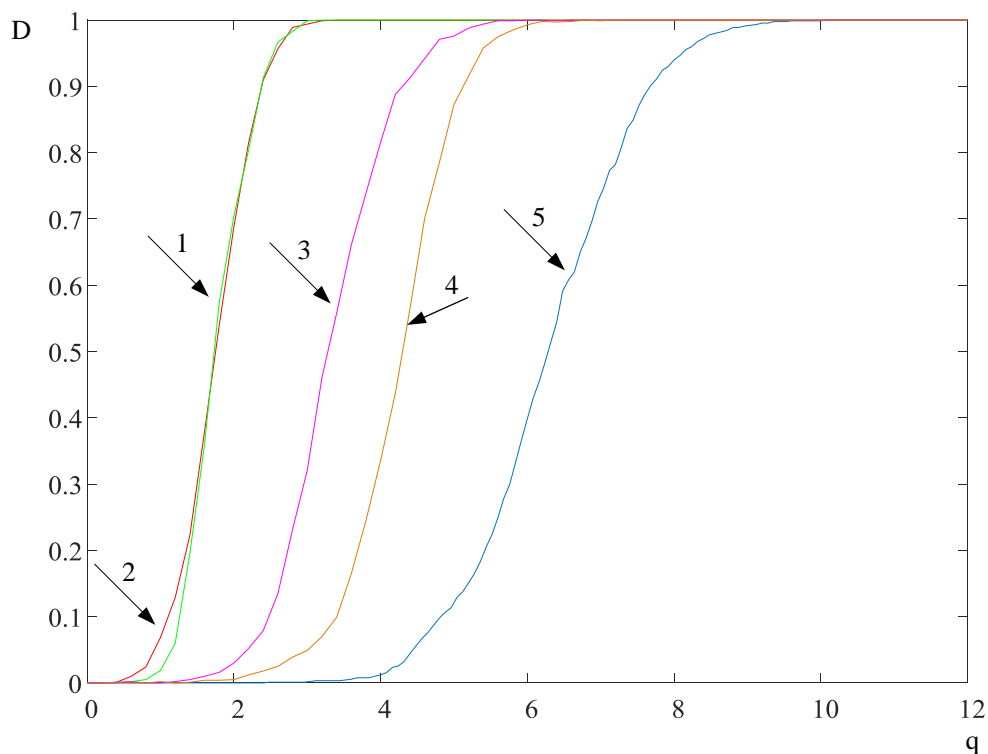
**Рис. 3. Количество элементов БПФ по размерности доплеровских частот, занимаемые распределенной целью в зависимости от угла направления траектории цели и расстояния до цели (длина цели 20 м.)**

### **Результаты математического моделирования корреляционного алгоритма обнаружения**

Методом математического моделирования были получены характеристики обнаружения целей корреляционным алгоритмом обнаружения для следующих случаев:

1. Характеристика обнаружения распределенной цели (корреляция площадок результатов БПФ двух радаров размером 2x2);
2. Характеристика обнаружения распределенной цели (корреляция площадок результатов БПФ двух радаров размером 4x4);
3. Характеристика обнаружения точечной цели (корреляция площадок результатов БПФ двух радаров размером 2x2);
4. Характеристика обнаружения точечной цели (корреляция площадок результатов БПФ двух радаров размером 4x4);
5. Стандартный алгоритм обнаружения в одиночном радаре(CFAR).

Полученные графики представлены на рисунке 4 в соответствии с вышеописанным порядком. График показывает, что размер площадки мало влияет на вид характеристики обнаружения. Это можно объяснить тем, что обнаружение цели в нашем случае зависит только от значения локального максимума двумерного БПФ, то есть от первого этапа двухэтапного алгоритма обнаружения, второй же этап при этом стабилизирует уровень ложных тревог.



**Рис. 4. Характеристики обнаружения для различных случаев**

### **Выводы**

Результаты моделирования показали, что накопление по площадке результатов корреляции для распределенной цели позволяет обнаруживать цели с большей вероятностью правильного обнаружения. Также на графиках виден выигрыш и по сравнению со стандартным обнаружением одним радаром (кривая 5), что ещё раз подтверждает эффективность применения корреляционного метода для обнаружения слабых целей.

### **Библиографический список**

1. **Черняк, В.С.** Многопозиционная радиолокация. – М.: Радио и связь, 1993.– 416 с.;
2. **Смирнова, Д.М.** Обнаружение и измерение координат движущихся наземных объектов в многопозиционной просветной радиолокационной системе: Дис. кандидата техн. наук: 05.12.14 / Нижегород. гос. техн. унт. Н. Новгород, 2012.-161 с.

УДК 621.396.96

## АЛГОРИТМ ФОРМИРОВАНИЯ РАДИОИЗОБРАЖЕНИЯ МАЛОПОДВИЖНЫХ ОБЪЕКТОВ В НЕКОГЕРЕНТНОМ ПЕРЕДВИЖНОМ РАДИОЛОКАТОРЕ

А.Д. Михайлов

ORCID: 0000-0001-6428-7109 e-mail: MikhailovAD@nntu.ru

А.В. Мякинков

ORCID: 0000-0001-6952-4134 e-mail: redvillage@mail.ru

Нижегородский государственный технический университет им. Р.Е. Алексеева

Исследование посвящено разработке алгоритма цифровой обработки сигналов для некогерентного передвижного радиолокатора, осуществляющего формирование радиоизображения малоподвижных объектов с обнаружением точек, обладающих наибольшей радиояркостью, в условиях сложного или заведомо неизвестного закона сканирования сектора обзора. В статье определен ряд проблем, связанных с решением поставленной задачи, и описаны способы их преодоления. При перемещении радиолокатора за время сканирования сектора обзора меняется ракурс обзора объектов, что приводит к проблеме «размытия» кадра. Решение проблемы «размытия» кадра при формировании радиоизображения может быть решена без учёта закона сканирования за счет описанного способа организации алгоритма обработки сигналов. Выбор формата представления радиоизображения и рабочей системы координат позволяют существенно сократить вычислительную сложность алгоритма, а также решить проблему совмещения кадров. Для анализа эффективности разработанных решений выполнено математическое моделирование, результаты которого приводятся в докладе.

**Ключевые слова:** радиоизображение, некогерентная обработка, обнаружение малоподвижных объектов, закон сканирования, система координат.

## AN ALGORITHM FOR GENERATING A RADIO IMAGE OF SEDENTARY OBJECTS IN AN INCOHERENT MOBILE RADAR

A.D. Mikhailov, A.V. Myakinkov

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

**Purpose:** The author provides an algorithm of digital signal processing for incoherent mobile radar, which generates a radio image of sedentary objects with the detection of points with the highest radio brightness in conditions of a complex or obviously unknown scanning law of the viewing sector.

**Design / methodology / approach:** The paper identifies a number of problems associated with solving the task and describes ways to overcome them. When the radar is moved during the scanning of the viewing sector, the aspect angle of objects changes, which leads to the problem of "blurring" the frame. The solution of the problem of "blurring" of the frame during the formation of a radio image can be solved without taking into account the scanning law due to the described method of organizing the signal processing algorithm. The choice of the radio image representation format and the operating coordinate system can significantly reduce the computational complexity of the algorithm, as well as solve the problem of framesfusion. To analyze the effectiveness of the developed solutions, mathematical modeling was performed, the results of which are presented in the paper.

**Findings:** The paper finds out optimal signal processing and data organization for incoherent mobile radar, which generates a radio image of sedentary objects with the detection of points with the highest radio brightness in conditions of a complex or obviously unknown scanning law of the viewing sector.

**Research limitations/implications:** Research is limited to incoherent mobile radar, which generates a radio image of sedentary objects. The developed algorithm is not available for mobile objects detection.

**Originality/value:** The developed algorithm is practically invariant to the scanning law that makes it universal for an arbitrary mechanical or electronic scanning technique. Thus, changing the construction of the radio frequency front-end, only the parametrization of the algorithm is needed.

**Key words:** radio image, incoherent processing, detection of sedentary objects, scanning law, coordinate system.

## Введение

В работе описываются некоторые технические решения, реализованные в алгоритме формирования радиоизображения для некогерентного передвигного радиолокатора, осуществляющего сбор данных о малоподвижных объектах. Данный радар имеет следующие особенности: радиолокатор является движущимся, а его целями являются малоподвижные или стационарные объекты, такие как ландшафт, здания и сооружения, любые объекты, скорость движения которых относительно земли пренебрежимо мала по сравнению с собственной скоростью перемещения радиолокатора; радиолокатор осуществляет сканирование пространства в пределах сектора обзора узким лучом; пакетом входных данных для алгоритма является вектор оценок мощности принимаемого сигнала, полученный при конкретном значении угловых координат луча (азимута и угла места в местной сферической системе координат радиолокатора). Закон сканирования, то есть траектория перемещения сканирующего луча в угловых координатах, может быть любым и заранее неизвестен. В радиолокаторе возможно применение исключительно некогерентных методов обработки информации. Требуется реализовать в алгоритме некогерентное накопление и обнаружение наиболее радиоярких точек.

Анализ источников показывает фактическое отсутствие строгого определения радиоизображения. Это могут быть данные, организованные в виде плоской матрицы [1,2]. В этом случае радиоизображение является двумерной проекцией окружающего пространства на сферу с центром в антенной решётке. Так же, данные могут быть представлены в произвольном формате, чаще всего являясь трёхмерным радиоизображением окружающего пространства [3,4,5]. В данной статье рассмотрим несколько способов обработки радиолокационной информации, тесно связанных с соответствующими форматами представления выходных данных, то есть радиоизображения.

## Кадровая обработка

Входные данные поступают к алгоритму в виде вектора, содержащего информацию об измеренной мощности принятого сигнала в зависимости от дальности при определённом положении сканирующего луча  $\mathbf{f}_{\alpha,\beta} = [(p_1, r_1), (p_2, r_2), \dots, (p_N, r_N)]$ . Этот вектор обрабатывается пороговым устройством. По результатам пороговой обработки может быть сформирована одна или несколько точек  $f_{\alpha,\beta, r_i} = p_i$ , соответствующих поверхности некоторого объекта (рис. 1).

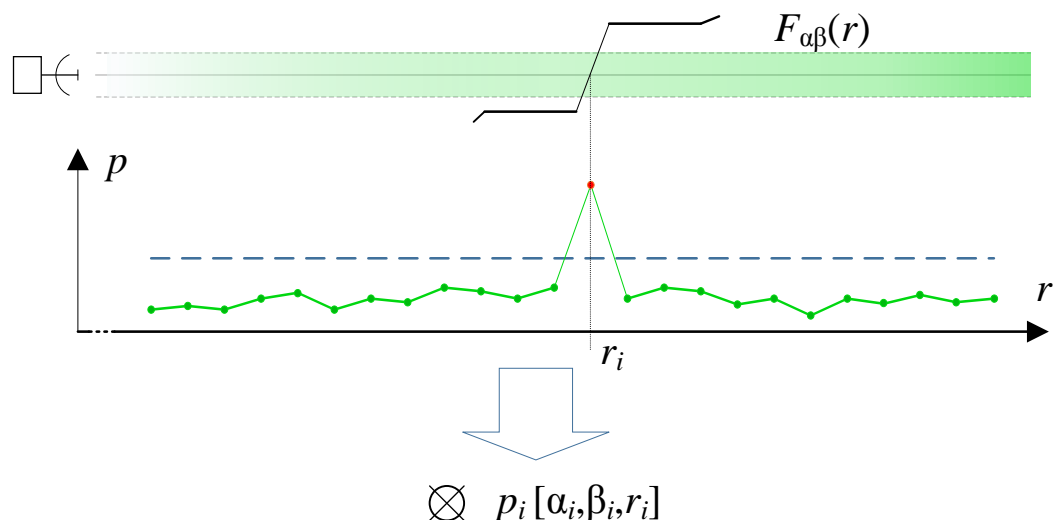


Рис. 1. Пороговая обработка входного пакета данных

Допустим, что за время сканирования сектора обзора формируется  $M$  векторов данных  $\mathbf{f}_{\alpha,\beta}$ , образуя некоторую объединённую матрицу данных  $\mathbf{P}_k(\alpha, \beta, r)$ . Тогда массив точек  $f$ ,

полученный после пороговой обработки матрицы данных можно считать плоским радиоизображением  $F_k(\alpha, \beta, r)$ , то есть проекцией объектов, попавших в сектор обзора радиолокатора. Каждое такое изображение будем называть кадром.

Если РЛС относительно целей неподвижна, ракурс обзора объектов не меняется, а значит кадры различаются между собой исключительно за счёт шумовой составляющей. С помощью сложения оценок мощностей точек кадров с совпадающими координатами  $\alpha, \beta$  и  $r$  обеспечивается межкадровое некогерентное накопление, тем самым, повышая отношение сигнал-шум.

$$F'(\alpha, \beta, r) = \sum_k F_k(\alpha, \beta, r) \quad (1)$$

При движении радиолокатора возможность применения формулы (1) оказывается ограничена. В случае, если скорость движения РЛС достаточна для существенного изменения ракурса обзора объектов, координаты ярких точек в кадрах не будут совпадать (рис. 2). Тогда перед сложением необходимо осуществлять экстраполяцию каждого предыдущего кадра к последнему и компенсировать эффект «размытия» кадров.

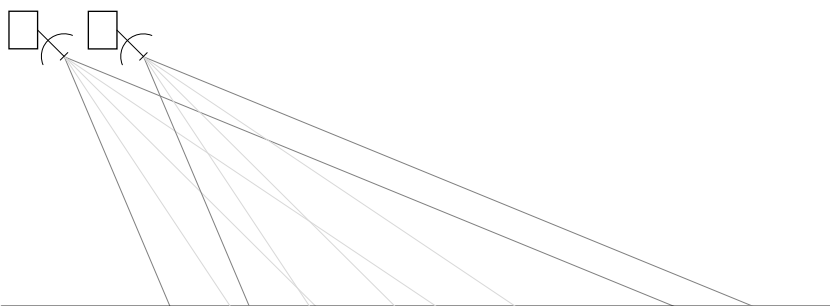


Рис. 2. Смещение сканируемой области при перемещении РЛС

### Эффект «размытия» кадра

Сканирование пространства в РЛС осуществляется непрерывно, то есть смена ракурса обзора объектов осуществляется непосредственно в процессе формирования кадра, что приводит к искажениям. Предположим, что сектор обзора сканируется лучом построчно, а РЛС движется равномерно и прямолинейно. Для этого случая эффект «размытия» кадра проиллюстрирован на рис. 3. В зависимости от скорости движения РЛС при постоянной скорости обзора, искажения кадра будут проявляться в разной степени.

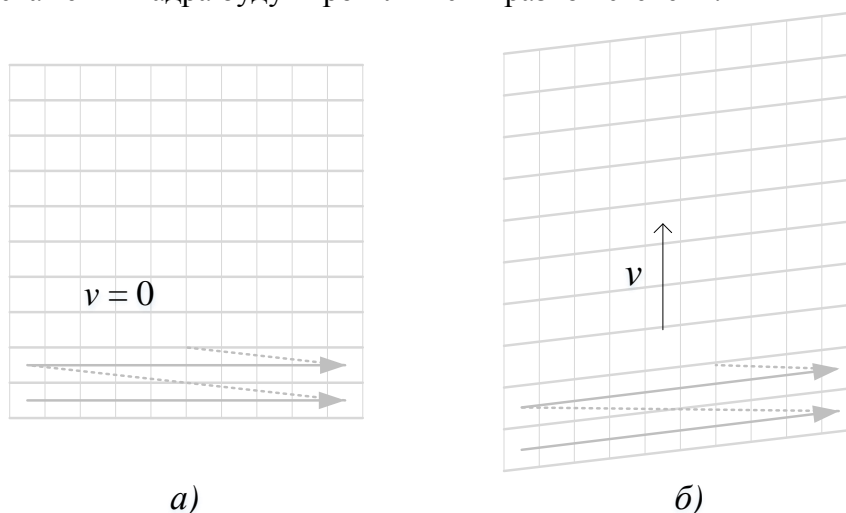


Рис. 3. Эффект «размытия» кадра

В случае построчного сканирования, проиллюстрированного на рисунке 3, а также прямолинейного равномерного движения РЛС, данный эффект может быть нивелирован.

Однако, закон сканирования, как и траектория движения радара, могут быть произвольными, из-за чего реализация покадровой обработки становится невозможна. Кроме того, стандартное матричное представление радиоизображения в вычислительной системе может быть эффективно только при законе сканирования с постоянным шагом и фиксированным числом дискретных значений по обеим осям угловых координат. Единственным форматом представления кадра, не зависящем от фактического закона сканирования, является вектор первичных отметок  $f_{\alpha,\beta,r_i} = p_i$ ; каждый элемент такого вектора содержит в себе информацию о трёх координатах отметки, а также непосредственно оценку мощности принятого сигнала. Индексы элементов такого вектора не имеют линейной связи с координатами отметок, а, следовательно, и с индексами отметок соседних кадров через пространственные координаты. Тогда, любое слияние кадров превращается в независимую обработку отдельных отметок  $f_{\alpha,\beta,r_i} = p_i$ .

### Поточечная обработка

Координаты отметки, полученной по результатам порогового анализа входного пакета данных, зависят от положения сканирующего луча в момент её получения, а значит, могут быть произвольными в некоторых пределах, обусловленных техническими возможностями РЛС: минимальная и максимальная дальность, крайние значения угловых координат сектора обзора. С учётом перемещения РЛС, диапазон возможных координат экстраполированных отметок ожидаемо расширяется. Совместная обработка должна проводиться между точками, для которых определено их пространственное соответствие. При этом одна из этих точек обязательно будет получена с обработки текущего входного пакета, а вторая – экстраполированная. Установление пространственного соответствия точек в таком случае должно определяться с учётом изменившегося ракурса обзора цели.

В качестве рабочей системы координат выбрана трёхмерная декартова система. Она позволяет выполнять экстраполяцию отметок с наименьшими затратами вычислительной мощности, а также позволяет избежать перерасчётов, связанных с изменением ракурса обзора, поскольку представляет представить массив точек не как двумерную проекцию, а как трёхмерное пространство. Каждая очередная отметка, полученная по результатам порогового анализа, пересчитывается в декартову систему координат. Далее, с каждой итерацией координаты пересчитываются (экстраполируются) с учётом перемещений РЛС (рис. 4).

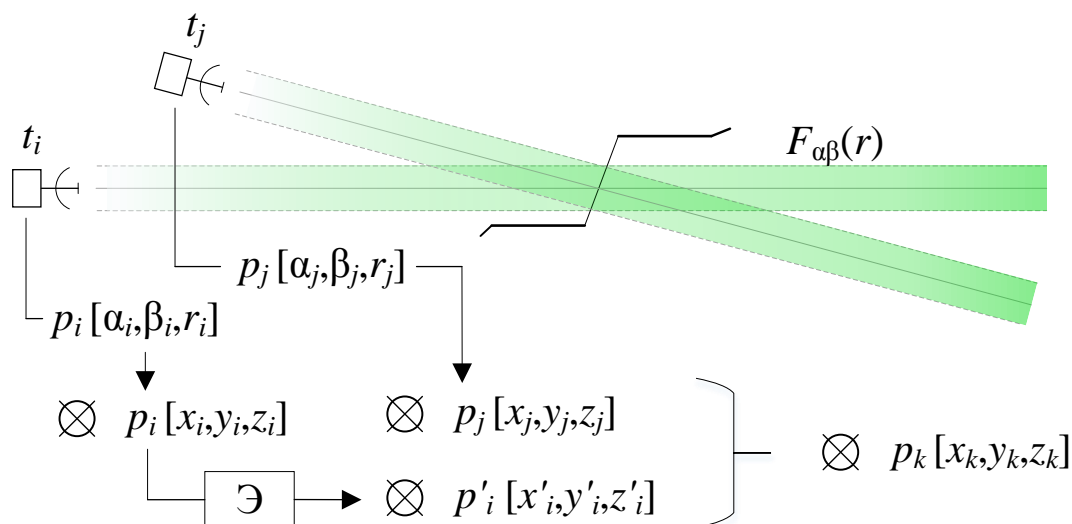


Рис. 4. Иллюстрация алгоритма поточечной обработки

В случае, если расстояние  $d$  между новой и экстраполированной отметкой окажется менее некоторого значения  $d_{min}$ , происходит «слияние» отметок. Координаты и оценка

мощности (яркость) новой отметки являются средним от соответствующих значений двух исходных.

$$[p_k, x_k, y_k, z_k]^T = \frac{[p'_i, x'_i, y'_i, z'_i]^T + [p'_j, x'_j, y'_j, z'_j]^T}{2} \quad (2)$$

Значение  $d_{min}$  определяется с учётом дальности отметки до РЛС  $r_i$  и заранее настраиваемого параметра  $C_d$ :  $d_{min} = C_d \cdot r_i$ . Настройка параметра осуществляется исходя из следующих зависимостей: его уменьшение приводит к повышению точности оценки координат до некоторого предельного значения, определяемого разрешающей способностью РЛС, однако, одновременно с этим уменьшает предельно допустимую собственную скорость цели, так как если за время смены ракурса обзора цель изменит своё местоположение на величину, превышающую  $d_{min}/2$ , две отметки не объединятся.

### Оптимизация алгоритма

Время работы РЛС не ограничено, тогда как число отметок, хранящихся в памяти, ограничено объёмом этой памяти. При этом, на каждой итерации алгоритма, совпадающей по времени с обработкой каждого отдельного пакета входных данных, требуется проводить экстраполяцию всех отметок, хранящихся в памяти, в соответствии с перемещением радиолокатора за данный промежуток времени. Основным объём располагаемых вычислительных ресурсов в этом случае будет задействован в экстраполяции.

В такой модели учитывается перемещение отметок относительно РЛС. Физически же именно РЛС перемещается относительно целей, тогда как сами цели неподвижны относительно друг друга. Это даёт возможность освободить значительный объём вычислительных ресурсов, отразив в модели неподвижность отметок. Для этого, начало рабочей системы координат необходимо закрепить на земле. При этом, РЛС окажется подвижна в данной системе координат (рис. 5). Эти перемещения отслеживаются так же, как и при экстраполяции отметок, однако вместо экстраполяции всех отметок в памяти требуется лишь пересчитать в рабочую систему координат новую отметку с учётом текущих координат и пространственной ориентации радара.

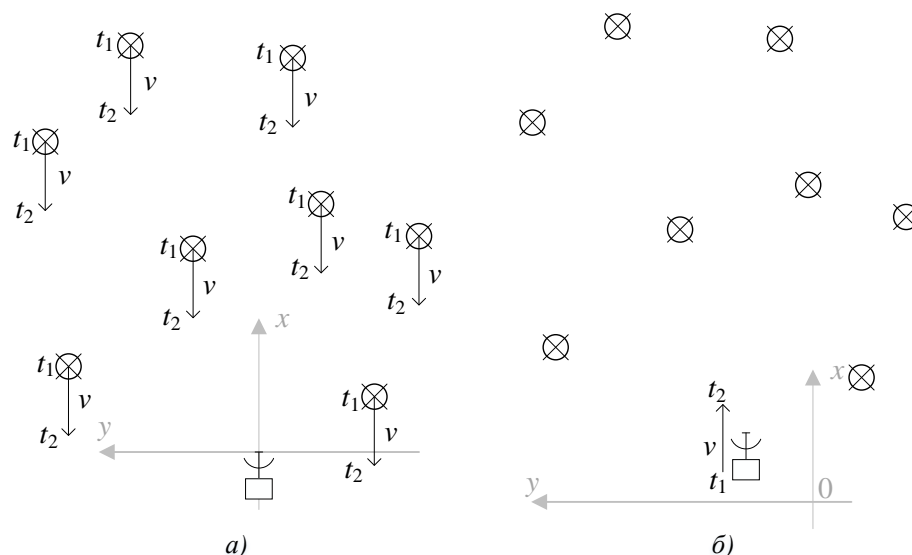
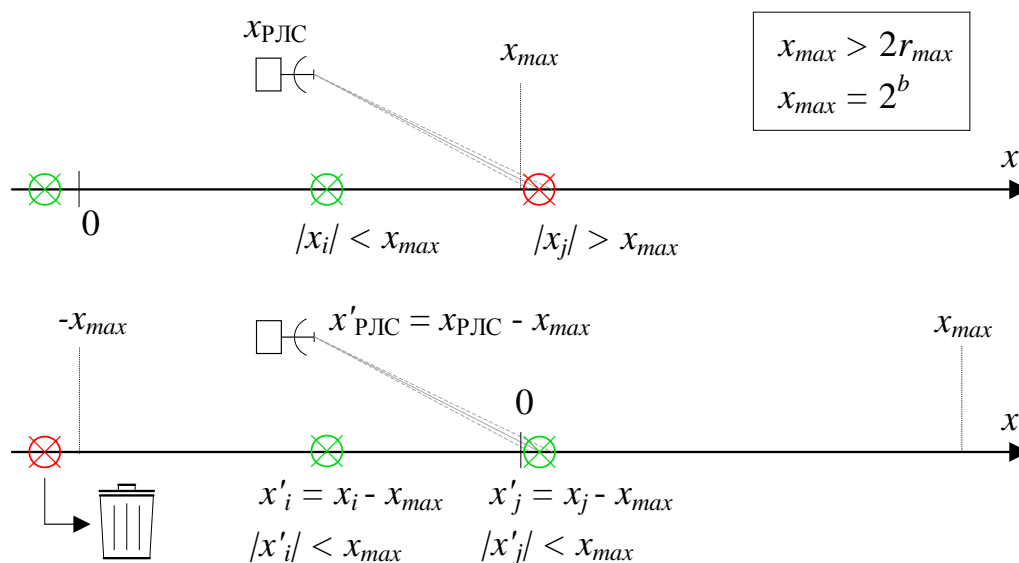


Рис. 5. Совокупность отметок в системе координат с: центром в РЛС (а), с центром, закреплённым на земной поверхности (б)

Расстояние, на которое РЛС может отдалиться от начала координат формально ничем не ограничено, то есть ничем не ограничены её координаты. Это создаёт угрозу



переполнения разрядной сетки по координатам. Решением данной проблемы является смещение начала координат. Рассмотрим алгоритм смещения на примере оси  $ox$  (рис.6).



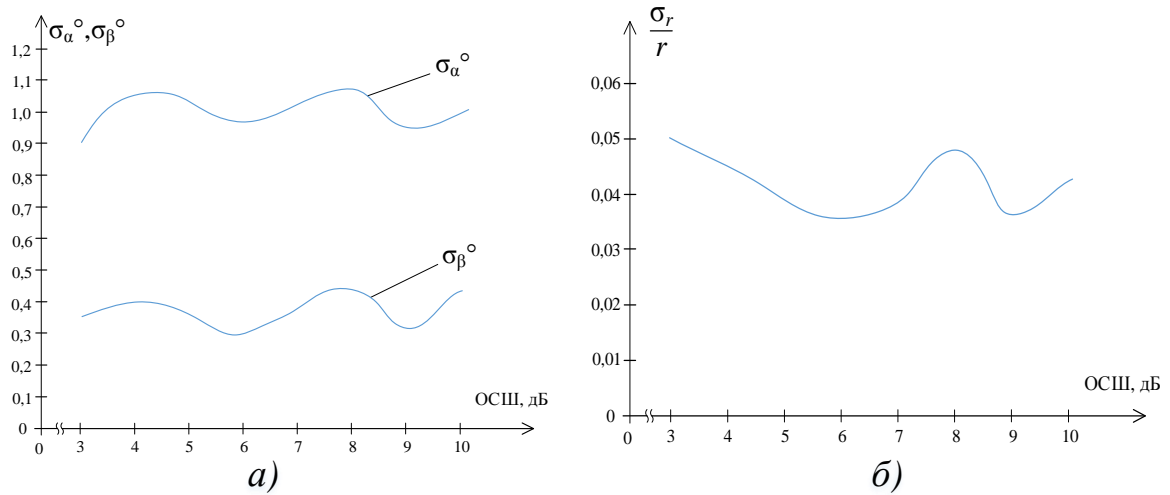
**Рис. 6. Иллюстрация алгоритма смещения координат**

Параметр  $x_{max}$  задаёт максимальное значение координаты по оси. Допустим, в очередной момент времени, алгоритм формирует отметку, координата которой вдоль оси превышает допустимое значение. Тогда, начало координат смещается вдоль этой оси на величину  $x_{max}$ . Технически, происходит пересчёт координат всех отметок и самой РЛС. Отметки, оказавшиеся после пересчёта за пределами области допустимых значений – удаляются. Величина  $x_{max}$  должна удовлетворять двум условиям: она должна быть больше двойной максимальной дальности сканирования, чтобы избежать слишком частого пересчёта  $x_{max} > 2r_{max}$ , параметр  $x_{max}$  следует выбрать равным степени двойки  $x_{max} = 2^b$ . В этом случае, операция пересчёта отметок станет значительно проще с точки зрения вычислительной сложности.

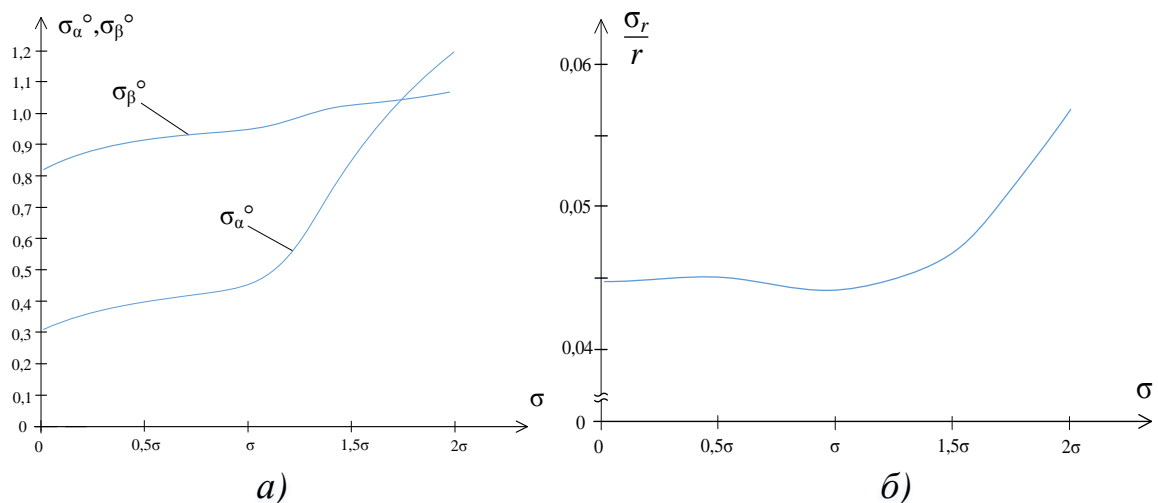
Совокупность неподвижных отметок, хранящихся в памяти устройства, представляют собой отображение наиболее ярких объектов в пространстве и является трёхмерным радиоизображением.

### Результаты моделирования

На основе данного алгоритма была построена программная модель для проверки точности оценки местоположения целей. Результаты представлены на рис. 7, 8. СКЗ ошибок измерения координат отметок близки к потенциальной точности их измерения (границе Рао-Крамера). Исходные данные были получены путём компьютерного моделирования работы РЛС по сканированию поверхности земли. На больших дальностях линейные размеры элемента разрешения по дальности существенно меньше, чем по углам, при этом, особенность исходных данных состояла в том, что угломестная координата сильно коррелировала с радиальной. Этим объясняется как большая точность определения угломестной координаты отметок в сравнении с азимутальной, так и большая чувствительность измерения азимута целей к СКО навигационных данных. Под навигационными данными понимаются данные по местоположению и ориентации в пространстве РЛС.



**Рис. 7. Зависимость СКЗ ошибки оценки координат от отношения сигнал/шум: СКЗ угловых координат (а), нормированного СКЗ радиальной координаты (б)**



**Рис. 8. Зависимость СКЗ ошибки оценки координат от СКО ошибки навигационных данных: СКЗ угловых координат (а), нормированного СКЗ радиальной координаты (б)**

### Заключение

По итогам исследования разработан алгоритм цифровой обработки сигналов для подвижной некогерентной РЛС. Алгоритм основан на известных принципах обработки радиолокационной информации. Однако особенности проектируемой системы, а также особенности назначения данной системы, а именно, обнаружение малоподвижных целей и построение их радиоизображения, позволили существенно оптимизировать обработку за счёт принятия некоторых неклассических технических решений, таких как неклассический выбор рабочей системы координат и определении вида и структуры радиоизображения.

Результаты моделирования подтверждают высокое качество обработки информации с точки зрения точности оценки их координат при существенной экономии вычислительных ресурсов системы.

### Библиографический список

1. **Crasto, N., Kishore Kumar, T., Anuradha, D., Baura, P., Nemani, S.** FPGA Implementation of Back Projection Algorithm for Radar Imaging/N.Crasto, T.Kishore Kumar, D. Anuradha, P. Baura, S. Nemani // 2013 International Conference on Radar. Adelaide, SA, Australia. – 2013.
2. **Yanik, M.E., Torlak, M.** Near-Field 2-D SAR Imaging by Millimeter-Wave Radar for Concealed Item Detection/M.E.Yanik, M.Torlak // 2019 IEEE Radio and Wireless Symposium (RWS). Orlando, FL, USA. – 2019.
3. **Millot, P., Castsnet, L., Casadebaig, L., Maaref, N., Gaugue, A.** An UWB Through-The-Wall radar with 3D imaging, detection and tracking capabilities/ P.Millot, L.Castsnet, L.Casadebaig, N.Maaref, A.Gaugue // 2015 European Radar Conference (EuRAD). Paris, France. – 2015.
4. **Geibig, T., Shoykhetbrod, A., Hommes, A., Herschel, R., Pohl, N.** Compact 3D imaging radar based on FMCW driven frequency-scanning antennas/T.Geibig, A.Shoykhetbrod, A. Hommes, R.Herschel, N.Pohl // 2016 IEEE Radar Conference (RadarConf). Philadelphia, PA, USA. – 2016.
5. **Gao, B., Zhang, F., Sun, G., Pan, S.** High-resolution 3D imaging with a photonics-based broadband MIMO radar/ B.Gao, F. Zhang, G. Sun, S. Pan // 2021 International Conference on Microwave and Millimeter Wave Technology (ICMMT). Nanjing, China.–2021.

## СХЕМА ПРОГРАММИРОВАНИЯ МЕМРИСТОРНЫХ ЭЛЕМЕНТОВ

Е.А. Букварев

ORCID: 0000-0003-0583-7662 e-mail: bukwarev@nntu.ru

К.С. Фомина

ORCID: 0000-0003-0531-209X e-mail: ksf96@mail.ru

Нижегородский государственный технический университет им. Р.Е. Алексева

С.А. Щаников

ORCID: 0000-0002-3938-8896 e-mail: seach@inbox.ru

Муромский институт (филиал) ФГУП высшего образования «Владимирский государственный университет имени Александра Григорьевича и Николая Григорьевича Столетовых»

Мемристор – полупроводниковый элемент, обладающий вольт-амперной характеристикой специального вида, содержащей линейный участок и две нелинейные области. Массивы мемристоров могут быть использованы в перспективных системах нейровычислений в качестве программируемого сопротивления (аналогового коэффициента умножения), при проведении операций аналогового умножения векторов, дискретного по времени. Для формирования требуемого сопротивления, мемристор должен быть подвергнут процедуре «программирования». Процедура может быть осуществлена двумя различными способами, связанными с различным направлением движения по вольт-амперной характеристике мемристора. В статье рассматривается предложенная схема универсального устройства программирования мемристора, выполненная на основе цифро-аналогового преобразователя и дискретного потенциометра с цифровым управлением. Показаны преимущества и недостатки различных способов программирования. Разработанный способ является универсальным, и не требует дополнительного анализа выходного напряжения.

**Ключевые слова:** мемристор, аналоговые нейровычисления дискретного времени, искусственный интеллект.

## MEMRISTOR ELEMENT PROGRAMMING CIRCUIT

E.A. Bukvarev, K.S. Fomina

NizhniyNovgorodstatetechnicaluniversityn.a. R.E. Alekseev

S.A. Shannikov

Murom Institute (branch) Federal state budgetary Educational Institution of Higher Education "Vladimir State University named after AlexanderGrigoryevich and NickolayGrigoryevichStoletovs"

**Purpose:**The article deals with the proposed memristor element programming circuit based on a digital-analog converter and a discrete potentiometer with digital control.

**Design/methodology/approach:**To form the required resistance, the memristor must be subjected to a "programming" procedure. The procedure can be carried out in two different ways, related to the different direction of movement according to the volt-ampere characteristic of the memristor. The article says about the advantages and disadvantages of various existing and possible programming methods. The developed method is universal and does not require additional analysis of the output voltage.

**Findings:**The proposed approach to the memristor element programming is based on a digital-analog converter and a discrete potentiometer with digital control.

**Research limitations/implications:**A memristor is a semiconductor element with a special type of volt-ampere characteristic containing a linear section and two nonlinear regions.

**Originality/value:**Arrays of memristors can be used in promising neural computing systems as a programmable resistance (analog multiplication ratio) when performing operations of analog multiplication of vectors, discrete in time.

**Keywords:** memristor, discrete time analog neuro processing, artificial intelligence.

## Вольт амперная характеристика мемристора

Вольт амперная характеристика (ВАХ) разработанных мемристоров показана на рис. 1. Допустимый ток не превышает  $\pm 0,5$  мА, диапазон напряжений находится в пределах  $-2..2$  В. Данный мемристор является достаточно низковольтным, и для его программирования во всех режимах достаточно, чтобы формирователи тока и напряжения работали в указанном диапазоне значений.

Для режима чтения с целью исключения деградации мемристора, напряжение источников сигнала  $V_1..V_n$  целесообразно ограничить диапазоном  $\pm 0,25$  В. Этот диапазон соответствует участку ВАХ, на котором отсутствует гистерезис.

При изменении напряжения на мемристоре, он последовательно проходит по ВАХ, как указано стрелками на рис. 1, а. Существует два граничных состояния мемристора, имеющих стандартные названия:

- LRS (LowResistanceState), низкоимпедансное состояние;
- HRS (High Resistance State), высокоимпедансное состояние.

По мере движения по ВАХ, мемристор может переходить из HRS в LRS. Такой переход называется «SET». Переход в обратном направлении, от LRS к HRS называется «RESET». На рис. 1, б, в показаны семейство экспериментально измеренных ВАХ с указанием переходов и статистика времени переключения между состояниями.

В [1] можно встретить принятое обозначения процедуры программирования в переходе «SET», именуемое «Decremental programming», а программирования в переходе «RESET» – «Incremental programming» соответственно. В названии в явном виде присутствует термин «уменьшение» в первом случае, и «увеличение» – во втором соответственно.

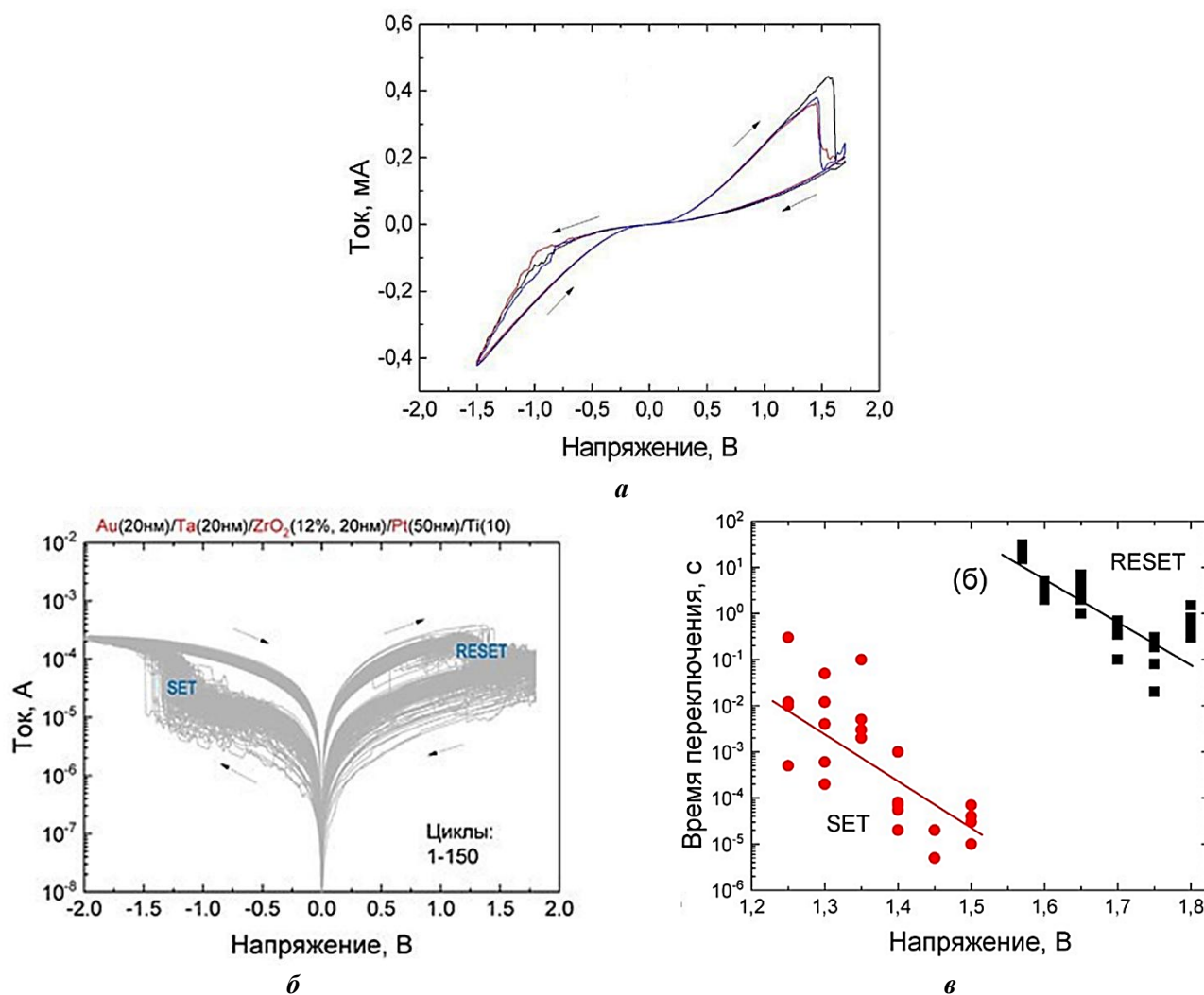


Рис. 1. Вольт-амперная характеристика: а) экспериментального образца мемристора кросс-бара; б) с указанием переходов; в) статистика времени переключения

### Переход в высокоимпедансное состояние

Типовая схема программирования в «RESET», заимствованная из [1], показана на рис. 2, а.

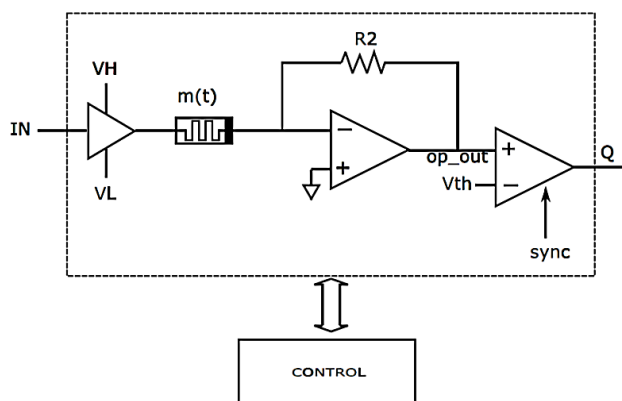
Подобные схемы программирования опробованы в [2], [3] и содержат источник программирующего напряжения, соединенный с программируемым мемристором  $m(t)$ , положительная клемма которого подается на вход преобразователя «ток-напряжение», выходное напряжение которого сравнивается с порогом  $V_{th}$  с помощью стробируемого сигналом «sync» компаратора.

Работает схема следующим образом. Вначале мемристор переводится в LRS состояние, затем на вход подаются импульсы длительностью 40 мкс и периодом повторения 100 мкс. С каждым новым импульсом проводимость мемристора постепенно уменьшается, а напряжение «op\_out» на выходе преобразователя увеличивается за счет уменьшения отношения  $R2/R(m(t))$ . Когда проводимость мемристора уменьшается до заданного значения, процесс прерывается, а мемристор считается запрограммированным.

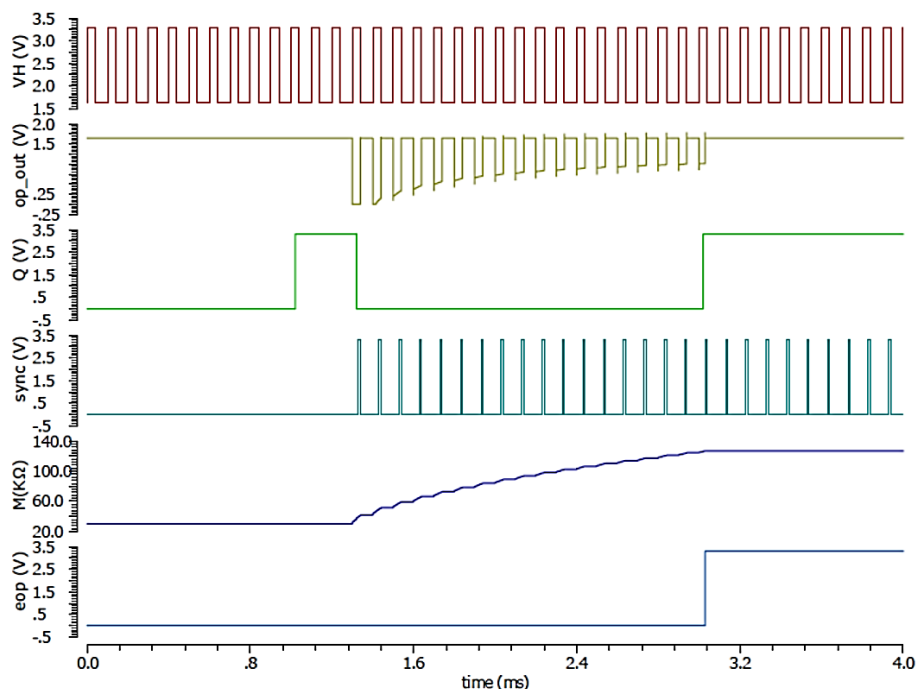
Диаграмма процесса показана на рис.2, б. Синим цветом (график  $M(\text{кОм})$ ) показано изменение сопротивления мемристора.

В целом, такой способ удобен тем, что его возможно реализовать в рамках совмещенной архитектуры управления кросс-баром, при которой источник рабочего напряжения может быть совмещен с источником напряжения программирования, приемник – с АЦП, а сам процесс программирования может быть распараллелен по строкам.

С другой стороны, характер изменения сопротивления мемристора в процессе программирования имеет достаточно нелинейный характер.



а



б

Рис. 2. «Incremental programming»: а) схема программирования; б) результаты моделирования схемы программирования

### Переход в низкоимпедансное состояние

Типовая схема программирования в «SET», заимствованная из [1], показана на рис. 3, а. Схема отличается от предыдущей тем, что программируемый мемристор располагается в цепи обратной связи преобразователя «ток-напряжение». Фактически, для мемристора операционный усилитель (ОУ) в такой схеме представляет собой источник тока величиной  $U_{in}/R1$ . Иными словами, схема программирования «SET» является токовой. Следом за преобразователем «ток-напряжение» расположен стробируемый сигналом «sync» компаратор с порогом  $V_{th}$ .

Диаграмма процесса программирования показана на рис.3, б.

В отличие от предыдущей схемы, здесь изменена полярность программирующих импульсов. Изначально мемристор переводится в HRS состояние, затем, по мере увеличения номера импульса, его сопротивление начинает уменьшаться. Синий график ( $M(\text{кОм})$ ) показывает изменение сопротивления мемристора. Процесс завершается, как только напряжение на выходе преобразователя «ток-напряжение» оказывается меньше порога  $V_{th}$ .

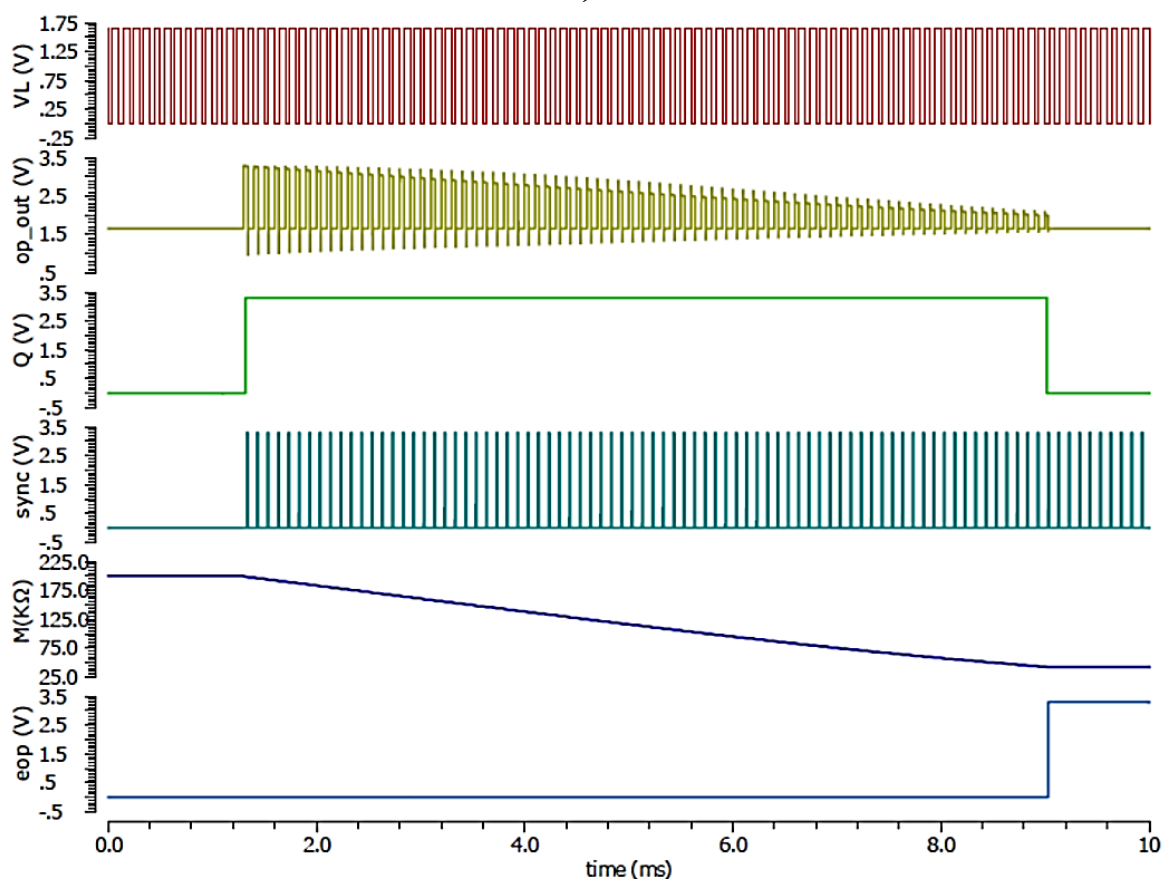
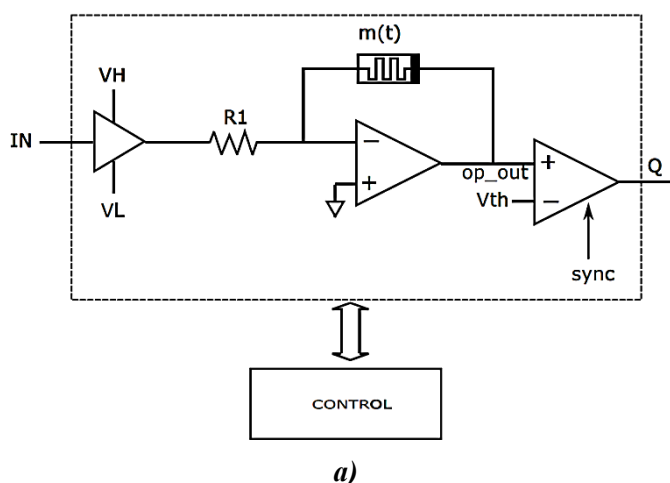


Рис. 3. «Decremental programming»: а) схема программирования; б) результаты моделирования схемы программирования

Способ программирования в процессе «SET» представляется более привлекательным по критерию лучшей линейности. Но при его реализации возникает необходимость коммутирования мемристора, что исключает возможность использования режима программирования, совмещенного с основным режимом работы устройства.

Что же касается ограничения тока мемристора для предотвращения выхода его из строя в процессе программирования, то способ «SET» оказывается удобнее по следующим причинам:

- ток мемристора жестко ограничен резистором R1 (рис. 3, а) в то время, как в схеме на рис. 2, а, при неудачном выборе ОУ, ток может оказаться ограничен лишь внутренним сопротивлением мемристора;

- процесс программирования начинается с малых токов, что является более безопасным.

### Схема программирования с использованием переменного резистора

На основе схемы с размещением мемристора в цепи обратной связи ОУ можно выполнить универсальное устройство программирования мемристора. На рис. 4 показана структурная схема подобного устройства. Она содержит биполярный формирователь напряжения программирования  $V_p$ , два идентичных канала ОУ DA1 и DA2, в одном из которых расположен программируемый мемристор MR, а в другом – образцовый переменный резистор RP, и компаратор DA3.

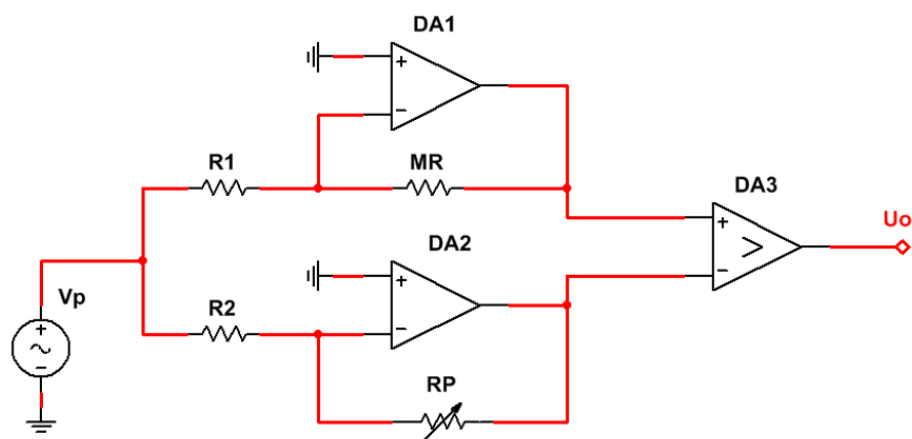


Рис. 4. Схема программирования с использованием переменного резистора

Процесс программирования начинается со сброса мемристора в состояние HRS. После чего переменный резистор программируется на нужное сопротивление, такое, какое должен будет иметь мемристор по окончании процесса программирования. Затем на выходе источника  $V_p$  формируется либо последовательность импульсов, подобно [1], либо линейно нарастающее напряжение. При этом в цепи DA1 через мемристор MR будет протекать постепенно возрастающий ток, а проводимость мемристора будет увеличиваться. В момент, когда внутреннее сопротивление MR окажется равным сопротивлению RP, напряжение на выходе DA1 превысит напряжение на выходе DA2, компаратор DA3 переключится, а напряжение  $U_o$  сменит полярность. Это будет являться критерием завершения процесса программирования.

Если источник напряжения  $V_p$  и ОУ выполнить двухполярными, тогда с помощью этой схемы можно осуществить программирование любого типа, как «Decremental», так и «Incremental», а также перевод мемристора в любое состояние без его перекоммутации.

На рис. 5 **Ошибка! Источник ссылки не найден.** показана работа программатора при значении сопротивления переменного резистора 50 кОм [4]. Выходной сигнал компаратора – « $V_{out}$ », показан синим цветом. В начальном состоянии оно равно логическому нулю. По мере уменьшения сопротивления мемристора, после в момент равенства его сопротивления сопротивлению переменного резистора, компаратор меняет логическое состояние, сигнализируя о завершении процедуры и необходимости снятия тока программирования.



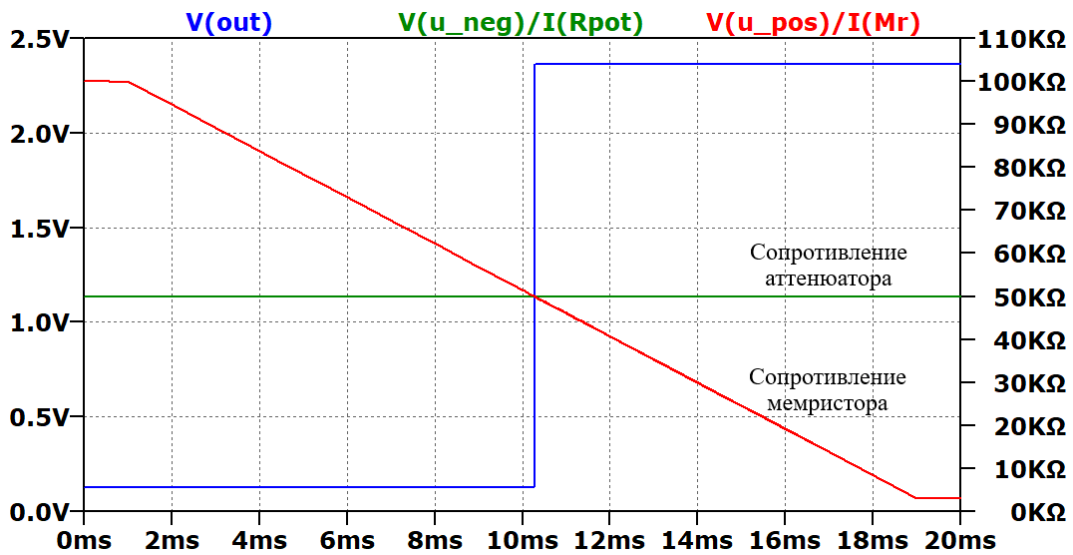


Рис.5. Результаты моделирования при сопротивлении аттенюатора 50 кОм

На рис. 6 показано семейство таких графиков при вариации сопротивления переменного резистора в диапазоне от 10 до 90 кОм с шагом 20 кОм. Моменты срабатывания компаратора оказываются точно в момент равенства сопротивлений мемристора и переменного резистора во всем диапазоне изменяющихся значений.

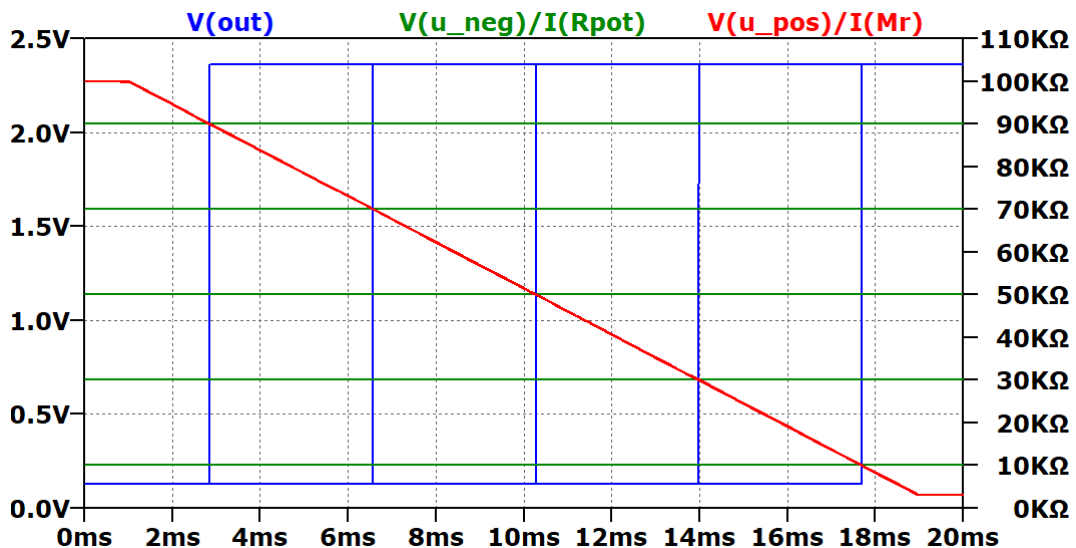


Рис. 6 Результаты моделирования при вариации сопротивления переменного резистора от 10 до 90 кОм

Модель показывает работоспособность заложенных в узел программатора структурных и схемотехнических решений.

### Выводы

Время сброса и программирования одного мемристора может достигать 0,05 с. Таким образом, процесс полного обновления матрицы размером 32×8, состоящей из 256 элементов, не будет превышать 15 с. К сожалению, это время определяется физическими особенностями мемристоров. Попытка распараллеливания процедуры программирования, например, по строкам кросс-бара приведет к возникновению токов перетока.

По результатам проведенного обзора, рациональным выбором представляется применение в составе устройства управления кросс-баром схемы программирования с использованием переменного резистора. Дополнительно к наличию компаратора в схему, показанную на рисунке 4 можно ввести и АЦП, вход которого будет подключен к выходу DA1. Таким образом, появляется возможность выбора средства измерения сопротивления мемристора как в процессе проведения программирования, так и для целей контроля сопротивления мемристора по окончании процедуры.

### Источник финансирования

Работа выполнена при поддержке Субсидии Министерства науки и высшего образования РФ (проект №13.2251.21.0098, соглашение №075-15- 2021-1017).

### Библиографический список

1. **O. A. Olumodeji, M. Gottardi** A pulse-based memristor programming circuit / Proc. IEEE International Symposium on Circuits and Systems (ISCAS) Conference Paper, 2017.
2. **L. Zhang, B. Li, X. Huang, C. Shen, S. Zhang, F. Burcea, H. Graeb, T.-Y. Ho, H. Li, U. Schlichtmann** An Efficient Programming Framework for Memristor-based Neuromorphic Computing Grace”, Design, Automation & Test in Europe Conference & Exhibition, Conference Paper, Publisher: IEEE, 2021.
3. **T.-W. Lee, J. H. Nickel** Memristor Resistance Modulation for Analog Applications”, IEEE Electron Device Letters, Volume: 33, Issue: 10, Journal Article, 2012.
4. **LTspice**. AnalogDevices.URL: <https://www.analog.com/en/design-center/design-tools-and-calculators/ltspice-simulator.html> (дата обращения 8.02.2022).

**ИМИТАЦИОННАЯ МОДЕЛЬ ИЗМЕРЕНИЙ ЭЛЕКТРОФИЗИЧЕСКИХ  
И ГЕОМЕТРИЧЕСКИХ ПАРАМЕТРОВ МЕТАМАТЕРИАЛОВ НА ОСНОВЕ  
SRR-ЭЛЕМЕНТОВ МЕТОДОМ ПОВЕРХНОСТНЫХ ЭЛЕКТРОМАГНИТНЫХ ВОЛН**

**А.И. Казьмин**

ORCID: 0000-0001-7682-2420 e-mail: [alek-kazmin@yandex.ru](mailto:alek-kazmin@yandex.ru)

**П.А. Федюнин**

ORCID: 0000-0001-6663-4362 e-mail: [fpa1969@yandex.ru](mailto:fpa1969@yandex.ru)

Военный учебно-научный центр Военно-воздушных сил «Военно-воздушная академия  
им. проф. Н. Е. Жуковского и Ю. А. Гагарина» (г. Воронеж)

Представлена имитационная модель, позволяющая оценивать точность измерения локальных значений электрофизических и геометрических параметров метаматериалов на основе SRR-элементов на металлической подложке методом поверхностных электромагнитных волн. Имитационная модель представляет собой совокупность функционально связанных между собой блока решения прямой задачи, блока минимизации целевой функции, блока имитирующего абсолютные погрешности измерений коэффициентов ослабления поля поверхностной электромагнитной волны и блока, реализующего определение коэффициентов ослабления поверхностной электромагнитной волны с помощью системы электродинамического моделирования. Блоки имитационной модели реализованы на основе системы электродинамического моделирования CSTMicrowavestudio и системы Matlab. Приведены результаты имитационного моделирования по измерению параметров метаматериала на основе прямоугольных SRR-элементов с областью отрицательной рефракции в диапазоне частот 10,06–10,64 ГГц. Численная проверка показала, что локальные значения эффективных электрофизических параметров исследуемого метаматериала отличаются от расчетных не более чем на 10 %.

**Ключевые слова:** метаматериал, имитационная модель, поверхностная электромагнитная волна, электрофизические параметры, обратная задача

**SIMULATION MODEL OF MEASUREMENTS OF ELECTROPHYSICAL  
AND GEOMETRIC PARAMETERS OF METAMATERIALS BASED ON  
SRR ELEMENTS BY THE METHOD OF SURFACE ELECTROMAGNETIC WAVES**

**A. I. Kaz'min, P. A. Fedyunin**

Zhukovsky and Gagarin Voronezh air force academy

**Purpose:** development of a simulation model that makes it possible to evaluate the accuracy of measuring local values of electrophysical and geometric parameters of metamaterial based on SRR elements by the method of surface electromagnetic waves.

**Design/methodology/approach:** the simulation model is a set of functionally interconnected block for solving the direct problem, a block for minimizing the objective function, a block that simulates the absolute errors in measuring the field attenuation coefficients of a surface electromagnetic wave, and a block that implements the determination of the attenuation coefficients of a surface electromagnetic wave using an electrodynamic modeling system. The simulation model is implemented on the basis of the CST Microwave studio electrodynamic modeling system and the Matlab system.

**Findings:** the simulation model of a real measuring complex has been developed, with the help of which, for various configurations of measuring complexes, it is possible to estimate the errors in determining the dielectric and magnetic permeability of various types of metamaterial. Simulation modeling showed that the local values of the effective electrophysical and geometric parameters of the studied metamaterial based on SRR elements differ from the calculated ones by no more than 10%.

**Research limitations/implications:** the present article provides a starting-point for further research in the measuring electrophysical parameters of metamaterial.

**Originality/value:** thus, it has been established that, in fact, the presented approach allows for local inspection of metamaterial parameters, which makes it possible to study how these parameters change over the surface of the sample under study and thereby carry out its quality control.

**Keywords:** metamaterial, simulation model, surface electromagnetic wave, electrophysical and geometric parameters, inverse problem

## Введение

Метаматериалы (ММ) благодаря своим уникальным свойствам (электрофизическим, радиофизическими и оптическим), которые не встречаются в природных условиях, нашли применение и открывают широкие перспективы по их применению в будущем в различных областях науки и техники [1–8].

Вместе с тенденцией развития ММ различных типов возрастают и требования по точности и достоверности к методам измерения их электрофизических и геометрических параметров (ЭФГП), к важнейшим из которых относятся частотные зависимости комплексных диэлектрической  $\dot{\epsilon}_{\text{ММ}}(f)$  и магнитной  $\dot{\mu}_{\text{ММ}}(f)$  проницаемостей, а также толщина ММ  $t_{\text{ММ}}$ .

Зачастую применение ММ в составе радиоэлектронных систем (антенны, подложки микрополосковых СВЧ-печатных плат и др.) подразумевает размещение их на металлической подложке. Исходя из этого, для эффективного применения ММ на практике в составе какого либо устройства, требуются достоверные данные по его ЭФГП с учетом размещения на металлической подложке.

С учетом размещения ММ на металлической подложке его ЭФГП начинают меняться и могут возникать новые резонансы в частотных зависимостях комплексных диэлектрической и магнитной проницаемостей. Исходя из этого, для их оценки требуются методы, обеспечивающие односторонний подход при измерениях. Методы свободного пространства, основанные на одновременном измерении комплексных коэффициентов отражения  $\dot{R}$  и прохождения  $\dot{T}$  электромагнитных волн, фактически оказываются непригодными для контроля ЭФГП ММ, нанесенных на металлическую подложку [9–15].

В [16] на основе исследования свойств ММ на основе прямоугольных SRR-элементов впервые представлен метод измерения ЭФГП ММ с помощью поверхностных электромагнитных волн (ПЭМВ) сверхвысокочастотного диапазона (СВЧ) диапазона. Метод основан на определении вектора ЭФГП ММ  $\xi_{\text{ММ}} = \{\epsilon'_{\text{ММ}}(f_k), \epsilon''_{\text{ММ}}(f_k), \mu'_{\text{ММ}}(f_k), \mu''_{\text{ММ}}(f_k), t_{\text{ММ}}\}$  по частотной зависимости комплексного коэффициента ослабления (КО) поля ПЭМВ, возбуждаемой в исследуемом образце:

$$\mathbf{A}_{\text{ММ}} = [\alpha'(f_k, \xi_{\text{ММ}}), \dots, \alpha'(f_K, \xi_{\text{ММ}})], \quad k = 1, \dots, K, \quad (1)$$

где  $\alpha'(f_k, \xi_{\text{ММ}}) = 1/J \sum_{j=1}^J \frac{1}{s} \ln(E_j / E_{j+1})$  – действительная часть КО поля ПЭМВ ММ,  $E_j$  и  $E_{j+1}$  – значения напряженности электрического поля ПЭМВ измеренные по нормали к поверхности ММ в точках измерений  $y$  и  $y + s$ ,  $s$  – расстояние между точками измерений,  $J$  – количество точек измерения,  $\xi_{\text{ММ}} = \{\epsilon'_{\text{ММ}}(f_k), \epsilon''_{\text{ММ}}(f_k), \mu'_{\text{ММ}}(f_k), \mu''_{\text{ММ}}(f_k), t_{\text{ММ}}\}$  – вектор ЭФГП ММ, где  $\epsilon'_{\text{ММ}}(f_k), \mu'_{\text{ММ}}(f_k)$  действительные и  $\epsilon''_{\text{ММ}}(f_k), \mu''_{\text{ММ}}(f_k)$  мнимые части комплексных диэлектрической и магнитной проницаемостей ММ соответственно,  $t_{\text{ММ}}$  – толщина ММ;  $K$  – количество частот измерений, путем решения операторного уравнения первого рода [16]:

$$\Theta_{\text{ММ}} \cdot \xi_{\text{ММ}} = \dot{\alpha}(f_k, \xi_{\text{ММ}}), \quad \xi_{\text{ММ}} \in \Xi_{\text{ММ}}, \quad \dot{\alpha}(f_k, \xi_{\text{ММ}}) \in \mathbf{A}_{\text{ММ}} \quad (2)$$

где  $\Xi_{\text{ММ}}$  – метрическое пространство векторов  $\xi_{\text{ММ}}$  для ММ,  $\mathbf{A}_{\text{ММ}}$  – метрическое пространство КО поля ПЭМВ для ММ;  $\Theta_{\text{ММ}}$  – нелинейный оператор для ММ, осуществляющий преобразование множества значений вектора  $\xi_{\text{ММ}}$  в множество значений КО поля ПЭМВ  $\dot{\alpha}(f_k, \xi_{\text{ММ}})$ .

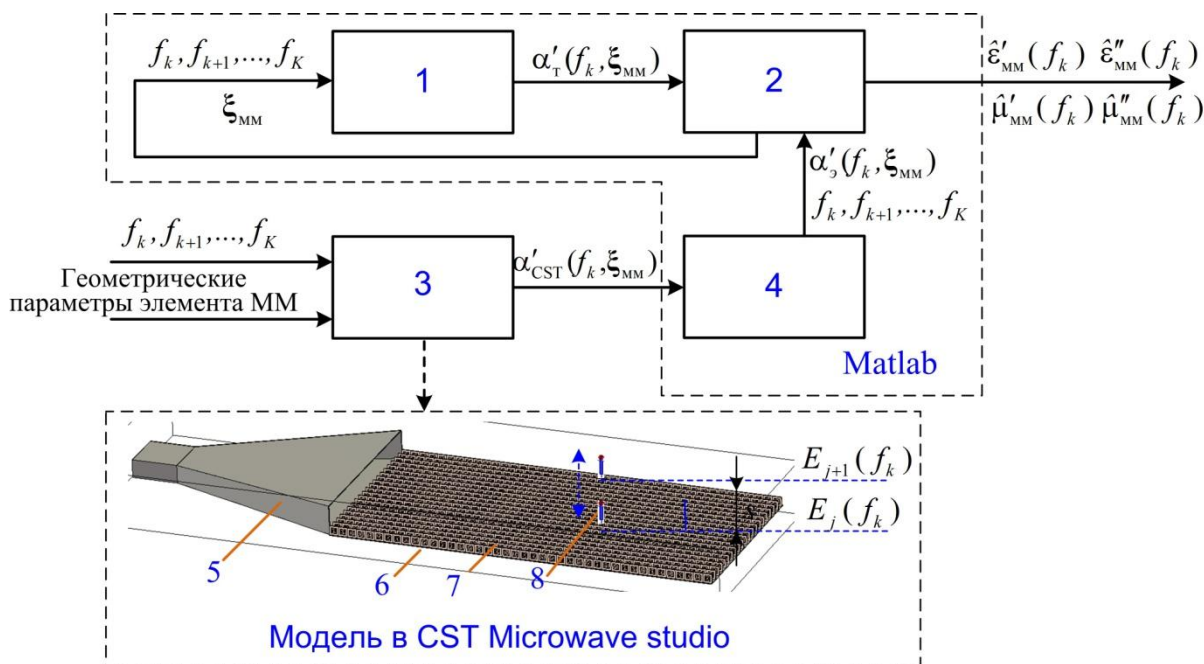
В направлении дальнейшего развития метода требуется более детально оценить его возможности по точности и достоверности оценки ЭФГП ММ с учетом размещения их на металлической подложке. Для более гибкой и оперативной оценки потенциальных возможностей метода в определении ЭФГП различных типов ММ целесообразно разработать имитационную модель измерений.

Целью статьи является разработка имитационной модели, позволяющей оценивать точность измерения локальных значений ЭФГП различных типов плоскостойких образцов метаматериалов на металлической подложке методом ПЭМВ [16].

### Имитационная модель измерений ЭФГП метаматериалов

Для более гибкой оценки точности и достоверности определения ЭФГП различных типов ММ разработана имитационная модель реального измерительного комплекса [16–17]. Структурная схема имитационной модели приведена на рис. 1, где обозначено: 1 – блок решения прямой задачи, 2 – блок решения операторного уравнения, 3 – блок формирования КО в системе CSTMicrowavestudio, 4 – блок формирования погрешностей измерений КО, 5 – антенная возбуждения ПЭМВ, 6 – металлическая подложка, 7 – исследуемый образец ММ, 8 – приемная антенна (измерительный зонд).

Такие элементы измерительного комплекса, как антенна возбуждения ПЭМВ 5, приёмная антенна (измерительный зонд) 8, а также исследуемый ММ 7 на металлической подложке 6, реализованы в блоке формирования КО в системе CSTMicrowavestudio. Геометрические размеры данных элементов точно соответствуют их размерам в реальном измерительном комплексе.



1 – блок решения прямой задачи, 2 – блок решения операторного уравнения, 3 – блок формирования КО в системе CST Microwave studio; 4 – блок формирования погрешностей измерений КО, 5 – антенна возбуждения ПЭМВ, 6 – металлическая подложка, 7 – исследуемый образец ММ, 8 – приемная антенна (измерительный зонд)

**Рис. 1. Имитационная модель измерений ЭФГП ММ методом поверхностных электромагнитных волн**

Блок определения теоретических значений комплексных коэффициентов ослабления 1 реализует вычисление левой части операторного уравнения (2), а блок решения операторного уравнения 2 реализует операцию по решению уравнения (2), путём варьирования составляющих вектора  $\xi_{MM}$ . Вывод в явном виде уравнения (2) и порядок его решения приведен в [16]. Данные блоки реализованы в системе Matlab. При этом решение обратной задачи (2) в блоке 2 осуществляется с помощью генетического алгоритма, реализованного в виде функции «ga» приложения GlobalSearch пакета Matlab.

Экспериментальные значения КО в имитационной модели (см. рис. 1) формировали в блоке формирования погрешностей измерений КО 4 в виде суммы значений модельных КО  $\alpha'_{CST}(f_k, \xi_{MM})$  и функции, имитирующей влияние шума. На точность определения коэффициентов ослабления

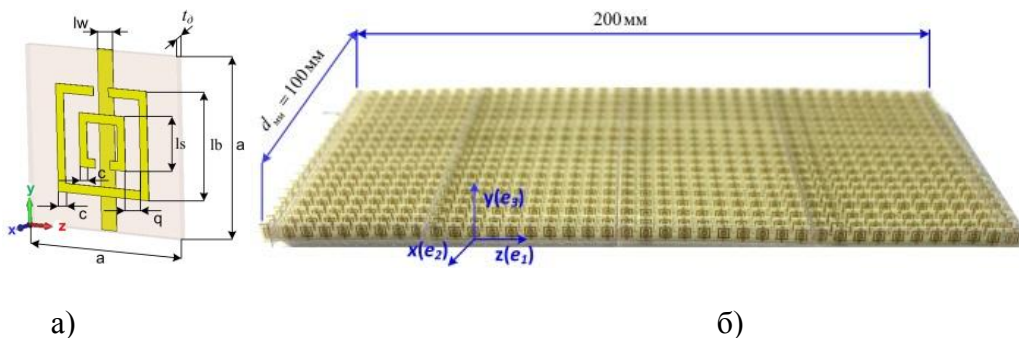
ПЭМВ влияет множество случайных факторов. Поэтому погрешность оценки указанных коэффициентов моделировалась аддитивным некоррелированным гауссовым шумом с нулевым математическим ожиданием  $m[n(f_k)] = 0$  и заданным уровнем дисперсии  $\sigma_k^2$  [16, 17].

Таким образом, экспериментальные значения КО ПЭМВ  $\alpha'_3(f_k, \xi_{MM})$  в блоке 4 (см. рис. 1) формировали на основе выражения [16–17]:

$$\alpha'_3(f_k, \xi_{MM}) = \alpha'_{CST}(f_k, \xi_{MM}) + n(f_k), k=1, 2, \dots, K$$

где  $n(f_k)$  – шум, действующий на k-измерение и обусловленный инструментальными погрешностями измерительного комплекса и другими погрешностями при проведении эксперимента.

Для численной проверки исследован образец плоскостойкого ММ на основе SRR-элементов с областью отрицательной рефракции в диапазоне частот 10,06–10,64 ГГц. Единичный элемент был спроектирован в системе CSTMicrowavestudio по методике приведенной в [15]. Внешний вид элемента приведен на рис. 2а, а геометрические размеры в табл. 1. С учетом геометрических параметров полученного единичного элемента изготовили образец ММ в виде плоской пластины внешний вид которого приведен на рис. 2б.



а) – единичный элемент ММ, б) – единичный элемент в составе плоскостойкого образца ММ

**Рис. 2 Исследуемый образец ММ**

**Таблица 1.**

**Геометрические размеры единичного элемента ММ, мм**

Параметр	Оптимизированное значение, мм
a	5
lw	0,5
ls	1,5
lb	3
c	0,25
q	0,5
t <sub>d</sub>	0,254

Не привязываясь строго к теории гомогенизации, для обеспечения проявления ММ свойств однородной среды, SRR-элементы в направлениях  $e_1$  и  $e_2$  (рис. 2) разместили с периодичностью 5 мм. Толщина ММ соответствует 1 элементу (5мм). Экспериментальное исследование образца ММ приведено в [16]. Исходя из цели статьи более подробно остановимся на численном эксперименте по определению ЭФГП данного образца ММ на разработанной имитационной модели.

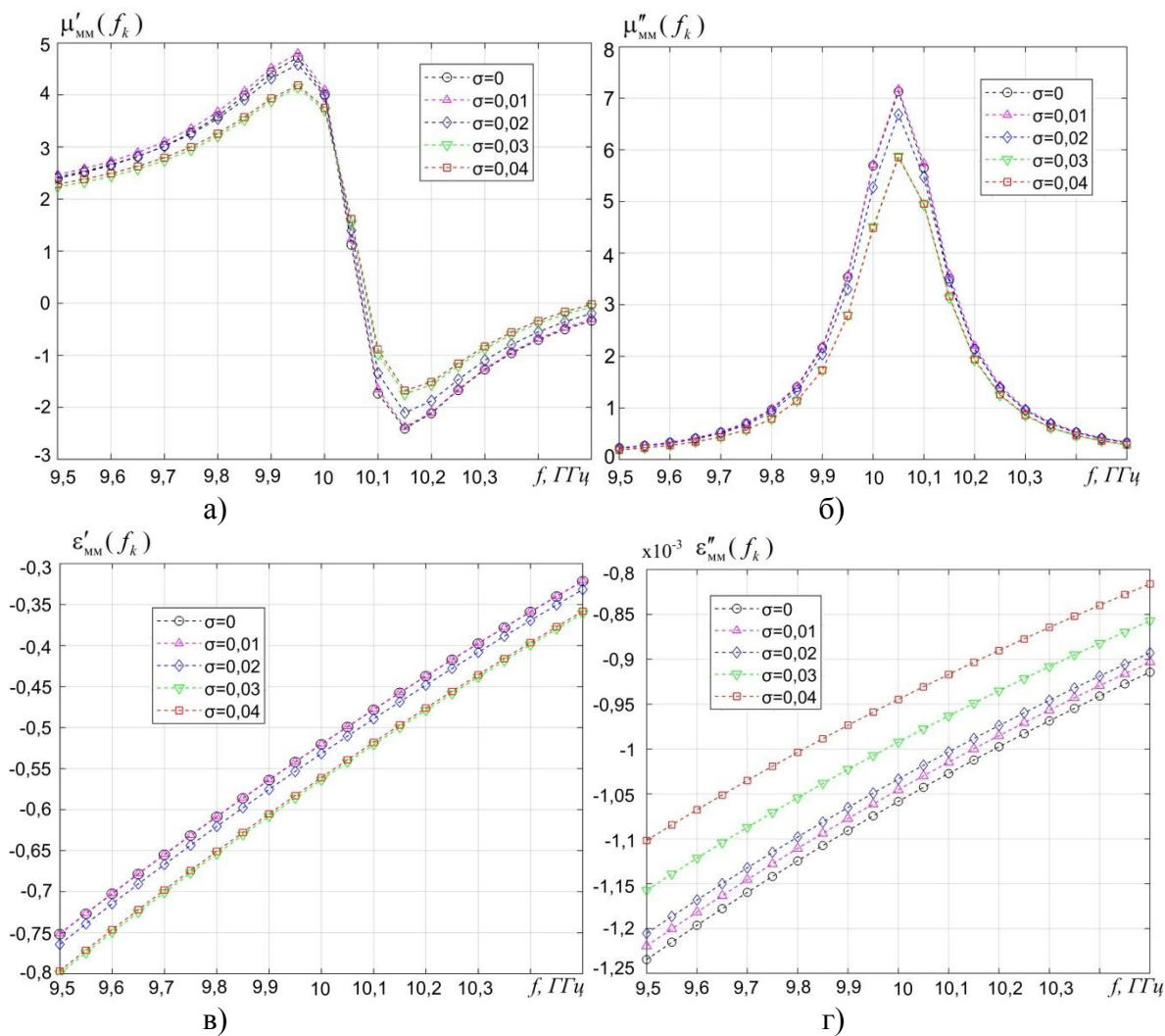
### Результаты численных экспериментов и их обсуждение

Для решения операторного уравнения (2) использовали 21 значение частоты из интервала  $\Delta f = 9,5 - 10,5$  ГГц с шагом по частоте 0,25 ГГц.

На рис. 3 приведены частотные дисперсионные зависимости комплексных диэлектрической и магнитной проницаемости полученные путем решения уравнения (2) при возрастающих значениях уровня среднего квадратического отклонения (СКО) шума в измерительных данных ( $\sigma=0,01$ ,  $\sigma=0,02$ ,  $\sigma=0,03$ ,  $\sigma=0,04$ ).

Анализ зависимостей на рис. 3 показал, что в целом погрешности оценок действительной и мнимой частей комплексных диэлектрической и магнитной проницаемостей ММ, в том числе и в резонансной области, при СКО уровня шума 0,02 не превышают 10 %. Погрешности оценки толщины при этом получили следующие:  $\sigma=0,01$ ,  $\delta t_{\text{ММ}} \approx 3,5\%$ ;  $\sigma=0,02$ ,  $\delta t_{\text{ММ}} \approx 6\%$ ;  $\sigma=0,03$ ,  $\delta t_{\text{ММ}} \approx 10\%$ ;  $\sigma=0,04$ ,  $\delta t_{\text{ММ}} \approx 19\%$ .

Таким образом, в целом, проведенный численный эксперимент показал, что разработанная электродинамическая модель измерений обеспечивает устойчивое решение обратной задачи (2) при воздействии шума.



- а) – действительная часть комплексной магнитной проницаемости,
- б) – мнимая часть комплексной магнитной проницаемости,
- в) – действительная часть комплексной диэлектрической проницаемости,
- г) – мнимая часть комплексной диэлектрической проницаемости

**Рис. 3. Частотные зависимости действительной и мнимой частей диэлектрической и магнитной проницаемостей ММ при возрастании СКО шума**

### Выводы

Таким образом, представлены результаты исследования показавшие высокую эффективность по измерению ЭФГП ММ с помощью ПЭМВ СВЧ-диапазона. Разработанный



метод обеспечивает локальный контроль ЭФГП плоскостойких ММ на металлической подложке при одностороннем доступе. Разработана имитационная модель реального измерительного комплекса, с помощью которой для различных конфигураций измерительных комплексов можно оценить погрешности определения диэлектрической и магнитной проницаемостей различных типов ММ. Имитационное моделирование показало, что локальные значения эффективных ЭФГП исследуемого ММ на основе SRR-элементов отличаются от расчетных не более чем на 10 %.

### Библиографический список

1. **Веселаго, В.Г.** Электродинамика веществ с одновременно отрицательными значениями  $\epsilon$  и  $\mu$  / В.Г. Веселаго // Письма в ЖТФ. 2016. Том 42. Вып. 5. С. 65-72.
2. **Лагарьков, А.Н.** Электрофизика и электродинамика метаматериалов / А.Н. Лагарьков, В.Н. Кисель, А.К. Сарычев, В.Н. Семенов // Теплофизика высоких температур. 2010. Т. 48. № 6. С. 1031–1048.
3. **Лагарьков, А.Н.** Электрофизика и электродинамика метаматериалов [Электронный ресурс] / А.Н.Лагарьков, В.Н.Кисель, А.К.Сарычев, В.Н.Семенов // Ин-т теоретической и прикладной электродинамики РАН: офиц. сайт. URL: [http://www.itae.ru/science/topics/№1%20\(метаматериалы\).pdf](http://www.itae.ru/science/topics/№1%20(метаматериалы).pdf) (дата обращения: 03.05.2022).
4. **Вендик, И.Б.** Метаматериалы и их применение в технике сверхвысоких частот (Обзор) / И.Б.Вендик, О.Г.Вендик // Журнал технической физики. 2013. Т. 83. вып. 1. С. 3–28.
5. **Слюсар, В.** Метаматериалы в антенной технике: основные принципы и результаты / В. Слюсар // Первая миля. 2010. №3–4. С. 44–60.
6. **Балабуха, Н.П.** Эффект обратного излучения электромагнитных волн волноводной структурой из метаматериала / Н.П.Балабуха, А.А.Баширин, В.Н.Семенов // Письма в ЖЭТФ. 2009. том 89, вып. 10. С. 593–598.
7. **Митрохин, В.Н.** Экспериментальные исследования СВЧ устройств содержащих метаматериалы / В.Н.Митрохин, Д.С.Рыженко, В.А.Тягунов // Физика волновых процессов и радиотехнические системы. 2011. Том. 14. № 3. С. 43–53.
8. **Pendry, J.B.** Magnetism from conductors and enhanced nonlinear phenomena / J.B.Pendry, A.J.Holden, D.J.Robbins, W.J.Stewart// IEEE Transactions on Microwave Theory and Techniques. Nov 1999. V. 47. Iss. 11. P. 2075–2084. DOI: 10.1109/22.798002.
9. **Ran, L.** Experimental Study on Several Left-Handed Matamaterials / L.Ran, J. Huangfu, H.Chen, X. Zhang, K. Chen, T. Grzegorzcyk, J.Kong // Progress In Electromagnetics Research. 2005. V. 51. P. 249–279. DOI:10.2528/PIER04040502.
10. **Smith, D. R.** Determination of effective permittivity and permeability of metamaterials from reflection and transmission coefficients / D. R.Smith, S.Schultz, P.Markos, C. M.Soukoulis // Physical Review B. 65, 195104. Published 19 April 2002. DOI: <https://doi.org/10.1103/PhysRevB.65.195104>.
11. **Smith, D. R.** Electromagnetic parameter retrieval from inhomogeneous metamaterials / D. R.Smith, D. C.Vier, Th.Koschny, C. M.Soukoulis // Physical Review E. 71, 036617 – Published 22 March 2005. DOI: <https://doi.org/10.1103/PhysRevE.71.036617>.
12. **David, R. Smith** Calculation and measurement of bianisotropy in a split ring resonator metamaterial / David R. Smith, Jonah Gollub, Jack J. Mock, Willie J. Padilla, David Schuring// Journal of Applied Physics. 2006. 100, 024507. DOI: <https://doi.org/10.1063/1.2218033>.
13. **Sikder Sunbeam Islam** The Design and Analysis of a Novel Split-H-Shaped Metamaterial for Multi-Band Microwave Applications / Sikder Sunbeam Islam , Mohammad Rashed Iqbal Faruque, Mohammad Tariqul Islam// Materials. 2014. No. 7(7). P. 4994–5011. DOI: <https://doi.org/10.3390/ma7074994>.
14. **Simovski, C.** Backward Wave Region and Negative Material Parameters of a Structure Formed by Lattices of Wires and Split-Ring Resonators / C.Simovski, P.A.Belov, S.He // IEEE Transaction on Antennas and Propagation. Oct. 2003. V. 51. P. 2582–2345. DOI:10.1109 / TAP.2003.817554.



15. **Lubkowski, G.** Extraction of effective metamaterial parameters by parameter fitting of dispersive models / G.Lubkowski, R.Schuhmann, T.Weiland// Microwave and Optical Technology Letters. Feb 2007. V. 49. Iss. 2. P. 285–288. DOI: <https://doi.org/10.1002/mop.22105>.

16. **Казьмин, А.И.** Контроль электрофизических параметров метаматериалов методом поверхностных электромагнитных волн / А.И. Казьмин, П.А. Федюнин // Дефектоскопия. 2021. № 4. С. 51–67.

17. **Казьмин, А.И.** Многочастотный оптимизационный метод измерения частотных зависимостей электрофизических параметров диэлектрических и магнитодиэлектрических покрытий / А.И.Казьмин // Измерительная техника, 2021. № 9. С. 54–61.

**ТЕОРЕТИЧЕСКИЕ ОСНОВЫ ОПРЕДЕЛЕНИЯ ЭЛЕКТРОФИЗИЧЕСКИХ ПАРАМЕТРОВ МНОГОСЛОЙНЫХ МАТЕРИАЛОВ И ПОКРЫТИЙ С ПОМОЩЬЮ ПОВЕРХНОСТНЫХ ЭЛЕКТРОМАГНИТНЫХ ВОЛН**

**А.И. Казьмин**

ORCID: 0000-0001-7682-2420 e-mail: [alek-kazmin@yandex.ru](mailto:alek-kazmin@yandex.ru)

**П.А. Федюнин**

ORCID: 0000-0001-6663-4362 e-mail: [fpa1969@yandex.ru](mailto:fpa1969@yandex.ru)

Военный учебно-научный центр Военно-воздушных сил «Военно-воздушная академия им. проф. Н. Е. Жуковского и Ю. А. Гагарина» (г. Воронеж)

Представлено формализованное описание метода поверхностных электромагнитных волн, применяемого для определения электрофизических параметров многослойных материалов и покрытий. Разработанные принципы использования в измерительных целях поверхностных электромагнитных волн позволяют в рамках единого методического подхода определять электрофизические и геометрические параметры многослойных диэлектрических и магнетодиэлектрических материалов и покрытий с произвольным числом слоев и значениями электрофизических параметров слоев. Разработанные подходы позволяют в зависимости от типа слоев материала или покрытия, набора подлежащих оценке электрофизических параметров и точности их определения выбирать требуемый набор физически измеряемых характеристик поля поверхностных электромагнитных волн, ширину полосы частот измерений и их количество. Экспериментальные исследования двухслойных и трехслойных диэлектрических покрытий показали, что при ширине полосы частот измерений 9–13 ГГц обеспечивается оценка диэлектрической проницаемости слоев не больше 6 % и толщин слоев не больше 5,5 % с доверительной вероятностью 0,95.

**Ключевые слова:** поверхностные электромагнитные волны, измерение, многослойные материалы и покрытия, электрофизические и геометрические параметры, операторное уравнение, дисперсионное уравнение, целевая функция

**THEORETICAL FOUNDATIONS TO EVALUATE ELECTROPHYSICAL PARAMETERS OF MULTILAYER MATERIALS AND COATINGS USING SURFACE ELECTROMAGNETIC WAVES**

**A. I. Kaz'min, P. A. Fedyunin**

Zhukovsky and Gagarin Voronezh air force academy

**Purpose:** development on the basis of a systematic formalized description of the developed method of surface electromagnetic waves of general methodological principles for determining the electrophysical and geometrical parameters of various types of multilayer materials and coatings with a given accuracy and reliability.

**Design/methodology/approach:** the developed principles of using surface electromagnetic waves for measuring purposes make it possible, within the framework of a unified methodological approach, to determine the electrophysical and geometric parameters of multilayer dielectric and magnetodielectric materials and coatings with an arbitrary number of layers and values of the electrophysical and geometrical parameters of the layers.

**Findings:** the developed approaches make it possible, depending on the type of layers of a material or coating, the set of electrophysical parameters to be evaluated and the accuracy of their determination, to choose the required set of physically measured characteristics of the field of surface electromagnetic waves, the bandwidth of measurements and their number.

**Research limitations/implications:** the present article provides a starting-point for further research in the measuring electrophysical parameters of multilayer materials and coatings.

**Originality/value:** thus, it has been established that, in fact, the presented approach allows for local inspection of multilayer materials and coatings parameters, which makes it possible to study how these parameters change over the surface of the sample under study and thereby carry out its quality control.

**Keywords:** surface electromagnetic waves, measurement, multilayer materials and coatings, electrophysical and geometrical parameters, operator equation, dispersion equation, objective function

## Введение

Диэлектрические и магнитодиэлектрические материалы и покрытия (МП), предназначенные для работы в диапазоне сверхвысоких частот (СВЧ), широко востребованы во многих наукоемких сферах, в том числе микроэлектронной, аэрокосмической, машиностроительной и др.

Одним из способов достижения требуемых характеристик МП является реализация их в виде многослойных диэлектрических и магнитодиэлектрических структур с диэлектрическими и магнитными потерями [1–7]. При этом с точки зрения основного целевого назначения – работа в диапазоне СВЧ, эффективность многослойных МП будет определяться их электрофизическими и геометрическими параметрами (ЭФП), к важнейшим из которых относятся комплексные диэлектрическая и магнитная проницаемости  $\hat{\epsilon}$  и  $\hat{\mu}$ , а также толщина слоев МП  $t$  [1–11].

Радиоволновые методы измерения интегральных значений ЭФП однослойных плоскостойких МП в значительной степени отработаны. При этом определение ЭФП МП, имеющих многослойную структуру с диэлектрическими и магнитными потерями, значительно сложнее.

В [9–11] на основе выявленных фундаментальных закономерностей, связывающих ЭФП многослойных МП с характеристиками поля поверхностных электромагнитных волн (ПЭМВ), развиты методы, обеспечивающие достоверный контроль их ЭФП в процессе проектирования и испытаний на эффективность.

Между тем, реализовать свойства современных многослойных МП, возможно только при строгом соблюдении технологических режимов и допусков при их нанесении на реальном объекте, что возможно за счет разработки общих методологических принципов оценки ЭФП и параметров дефектов МП в едином цикле измерения в одних и тех же условиях (размещение на металлической подложке, конструкционном композиционном материале и др.).

Исходя из этого, основной целью статьи является разработка общих методологических принципов определения ЭФП многослойных МП с заданными точностью и достоверностью.

### Формализованное описание метода поверхностных электромагнитных волн

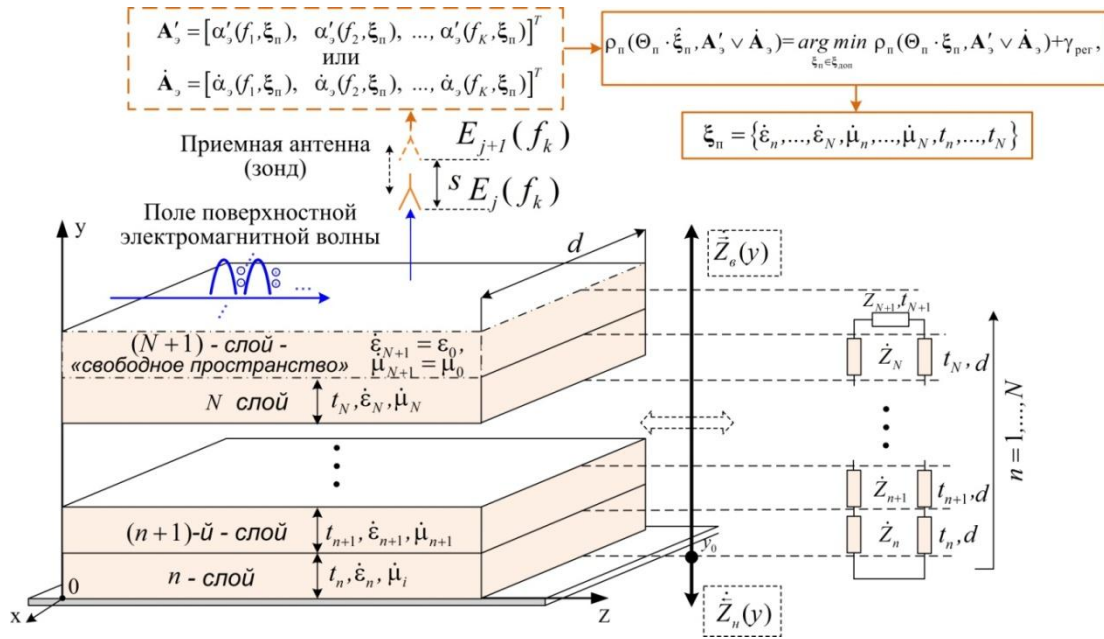
Основным информативным параметром ПЭМВ, используемым для проведения измерений ЭФП и дефектоскопии, выступает поперечное волновое число, характеризующее распределение поля ПЭМВ в свободном пространстве по нормали к поверхности МП – комплексный коэффициент ослабления (КО) поля ПЭМВ  $\hat{\alpha}(f_k, \xi_n) = \alpha'(f_k, \xi_n) - j\alpha''(f_k, \xi_n)$ , где  $\alpha'(f_k, \xi_n)$  – действительная часть КО ПЭМВ – коэффициент характеризующий экспоненциальное ослабление поля ПЭМВ по нормали к поверхности покрытия;  $\alpha''(f_k, \xi_n)$  – мнимая часть КО ПЭМВ – коэффициент характеризующий фазовый набег для ПЭМВ, распространяющейся по направлению к поверхности раздела «покрытие-свободное пространство»,  $\xi_n = \{\hat{\epsilon}_n, \dots, \hat{\epsilon}_N, \hat{\mu}_n, \dots, \hat{\mu}_N, t_n, \dots, t_N\}$  – вектор, подлежащих оценке ЭФП,  $\hat{\epsilon}_n, \dots, \hat{\epsilon}_N$  – комплексные диэлектрические проницаемости,  $\hat{\mu}_n, \dots, \hat{\mu}_N$  – комплексные магнитные проницаемости,  $t_n, \dots, t_N$  – толщины МП [9–11]. В [11–13] доказано, что действительную часть КО  $\alpha'(f_k, \xi)$  можно определить по результатам косвенных измерений напряженности электрического поля ПЭМВ по нормали к поверхности покрытия в разнесенных точках измерений методом «зонда»:

$$\alpha'(f_k, \xi_n) = \frac{1}{J} \sum_{j=1}^J \frac{1}{s} \ln \left[ \frac{E_j}{E_{j+1}} \right], \quad (1)$$

где  $E_j$  и  $E_{j+1}$  – значения напряженности электрического поля ПЭМВ, измеренные по нормали к поверхности покрытия в точках измерений  $y$  и  $y+s$ ,  $s$  – расстояние между точками измерений,  $J$  – количество точек измерения.

В общем случае для определения ЭФП МП требуется оценить КО в заданной полосе частот на сетке дискретных частот, т.е. получить частотную характеристику КО поля ПЭМВ

$\alpha'(f_k, \xi_n), f_k, k=1, 2, \dots, K$ . Представим задачу оценки ЭФГП многослойных МП в формализованном виде. Геометрия задачи измерения ЭФГП многослойных МП приведена на рис. 1.



- а) – геометрия задачи и обобщенная схема определения ЭФГП многослойных МП;  
 б) – эквивалентная схема многослойного МП на основе метода поперечного резонанса

**Рис. 1. Геометрия задачи и обобщенная схема определения ЭФГП многослойных МП на металлической подложке с помощью ПЭМВ**

Задача определения вектора ЭФГП и параметров дефектов  $\xi_n$  плоскостойких образцов многослойных МП по измеренным характеристикам поля ПЭМВ  $\mathbf{r}$  сводится к решению операторного уравнения первого рода:

$$\Theta_n \cdot \xi_n = \mathbf{r}, \quad \xi_n \in \Xi_n, \quad \mathbf{r} \in \mathbf{R}, \quad (2)$$

где  $\Xi_n$  – метрическое пространство векторов  $\xi_n$  для многослойного МП,  $\mathbf{R}$  – метрическое пространство характеристик поля ПЭМВ для многослойного МП;  $\Theta_n$  – нелинейный оператор для многослойного МП, осуществляющий преобразование множества значений вектора  $\xi_n$  в множество значений характеристик поля ПЭМВ  $\mathbf{r}$ .

Решением операторного уравнения (2) на компакте  $\Xi_n^k$  из  $\Xi_n$  будет являться элемент  $\xi_{0n} \in \Xi_n^k$ , минимизирующий функционал (целевую функцию)  $\rho_{nR}(\Theta_n \cdot \xi_n, \mathbf{r}_s)$ :

$$\rho_R(\Theta_n \cdot \xi_{0n}, \mathbf{r}_s) = \arg \min_{\xi_n \in \Xi_n^k} \rho_R(\Theta_n \cdot \xi_n, \mathbf{r}_s) + \gamma_{рег}, \quad (3)$$

где  $\mathbf{r}_s$  – полученные в эксперименте значения измеряемой характеристики поля ПЭМВ, то есть при наличии некоторой погрешности;  $\rho_R(\Theta_n \cdot \xi_n, \mathbf{r}_s)$  – расстояние между  $\Theta_n \cdot \xi_n$  и  $\mathbf{r}_s$  в метрическом пространстве  $\mathbf{R}$  возможных значений  $\mathbf{r}$ ;  $\gamma_{рег}$  – функция, регуляризирующая решение (2).

Область  $\Xi_n$  задания  $\xi_n$  определяется границами технически возможных параметров МП. Граничные значения области  $\mathbf{R}$  определяются максимальными значениями характеристик поля ПЭМВ.

Таким образом, для многослойного МП с границами технически возможных параметров  $\xi_{\min} \leq \xi_n \leq \xi_{\max}$ , необходимо найти вид нелинейного оператора  $\Theta_n$  и целевой функции  $\rho_R(\Theta_n \cdot \xi_{n0}, \mathbf{r}_s)$  операторного уравнения (2), обеспечивающих оценку неизвестных ЭФГП и параметров дефектов с заданными точностью и достоверностью при минимизации целевой функции, путем варьирования составляющими вектора  $\xi_n$ .

Для оценки качества многослойных МП определяется вектор ЭФГП  $\xi_n = \{\dot{\varepsilon}_n, \dots, \dot{\varepsilon}_N, \dot{\mu}_n, \dots, \dot{\mu}_N, t_n, \dots, t_N\}$ , где  $\dot{\varepsilon}_n, \dots, \dot{\varepsilon}_N$  и  $\dot{\mu}_n, \dots, \dot{\mu}_N$  – комплексные диэлектрические и магнитные проницаемости слоев,  $t_n, \dots, t_N$  – толщины слоев МП. В качестве физически измеряемых характеристик поля ПЭМВ  $\mathbf{r}$  выступает вектор-столбец ее комплексных коэффициентов ослабления (КО)  $\dot{\mathbf{A}}_s$  или вектор-столбец действительной части комплексного КО  $\mathbf{A}'$  [9–11].

Нелинейный оператор  $\Theta_n$  описывают распространение ПЭМВ в многослойном МП и представляют собой дисперсионное уравнение, составленное на основе метода «поперечного резонанса» [11, 12, 14].

### Математическая модель определения частотной зависимости комплексного КО

Модель представляет собой обобщенное дисперсионное уравнение для ПЭМВ *E*- или *H*-типов в многослойных МП, позволяющее однозначно определять частотную зависимость комплексного коэффициента ослабления ПЭМВ

$\dot{\mathbf{A}}_s = [\alpha'_{s1}(f_1, \xi_n) - j\alpha''_{s1}(f_1, \xi_n), \dots, \alpha'_{sK}(f_K, \xi_n) - j\alpha''_{sK}(f_K, \xi_n)]^T$ ,  $k = 1, \dots, K$  при задании вектора ЭФГП  $\xi_n = \{\dot{\varepsilon}_n, \dots, \dot{\varepsilon}_N, \dot{\mu}_n, \dots, \dot{\mu}_N, b_n, \dots, b_N\}$  и частоты  $f$ .

В качестве примера рассмотрим дисперсионное уравнение для ПЭМВ *E*-типа. В общем случае уравнение «поперечного резонанса» для составления дисперсионных уравнений для любых типов МП имеет следующий вид [11, 12, 14]:

$$\dot{Z}_b(y) + \dot{Z}_n(y) \text{ при } \forall y, \quad (4)$$

где  $\dot{Z}_b(y)$  и  $\dot{Z}_n(y)$  – эквивалентные характеристические сопротивления «вверх» и «вниз» (рис. 1) относительно произвольного опорного сечения  $y_0$  (для удобства математических преобразований выбрано сечение  $y_0$  между слоем покрытия и металлической подложкой).

Толщина типичных многослойных МП на практике обычно намного больше поверхностного слоя в металле, исходя из этого принято, что  $\dot{Z}_n(y) \approx 0$ . Характеристическое сопротивление  $\dot{Z}_b(y)$  «вверх» относительно опорного сечения  $y_0$  представляет собой эквивалентное характеристическое сопротивление слоев  $n = 1, \dots, N+1$  и определяется по рекуррентной формуле трансформации волновых сопротивлений [14].

Итоговое дисперсионное уравнение имеет следующий вид:

$$\Theta_1 \equiv D[\dot{\alpha}_t(f_k, \xi_n); f_k] \equiv \dot{Z}(n, \dots, N+1) = 0. \quad (5)$$

Полученное дисперсионное уравнение позволяет однозначно определять значения комплексных КО поля ПЭМВ  $\dot{\alpha}_t(f_k, \xi_n)$  при задании составляющих вектора ЭФГП  $\xi_n = \{\dot{\varepsilon}_n, \dots, \dot{\varepsilon}_N, \dot{\mu}_n, \dots, \dot{\mu}_N, b_n, \dots, b_N\}$  и частоты  $f_k$ .

### Целевая функция и регуляризация решения операторного уравнения

В соответствии с [8] решение операторного уравнения (2) и, соответственно, определение вектора ЭФГП  $\xi_n = \{\dot{\varepsilon}_n, \dots, \dot{\varepsilon}_N, \dot{\mu}_n, \dots, \dot{\mu}_N, b_n, \dots, b_N\}$  многослойных МП, сводится к минимизации

целевой функции, построенной, как невязка между вектор-столбцом экспериментальных  $\mathbf{A}_3$  и вектор-столбцом вычисленных теоретических значений  $\mathbf{A}_T$  комплексных КО ПЭМВ, на наборе частот  $f_k, k=1, 2, \dots, K$ , в заданной полосе частот  $\Delta f = f_k - f_1$ :

$$\hat{\xi}_n = \arg \min_{\xi_n \in \xi_{\text{доп}}} \rho(\Theta_n \cdot \xi_n, \mathbf{A}_3) = \frac{1}{K} |\mathbf{A}_3 - \mathbf{A}_T|^2 + \varpi \left( \sum_{n=1}^N q_n^{\dot{\epsilon}} \|(\dot{\epsilon}_n - \dot{\epsilon}_{n0})\|^2 + \sum_{n=1}^N q_n^{\dot{\mu}} \|(\dot{\mu}_n - \dot{\mu}_{n0})\|^2 + \sum_{n=1}^N q_n^t \|(t_n - t_{n0})\|^2 \right), \quad (6)$$

где  $\mathbf{A}_3$  и  $\mathbf{A}_T$  – вектор-столбцы, со следующими компонентами:

$$\mathbf{A}'_3 = [\alpha'_3(f_1) - j\alpha''(f_1, \Phi_\alpha), \alpha'_3(f_2) - j\alpha''(f_2, \Phi_\alpha), \dots, \alpha'_3(f_K) - j\alpha''(f_K, \Phi_\alpha)]^T;$$

$$\mathbf{A}'_T = [\alpha'_T(f_1, \xi_n), \alpha'_T(f_2, \xi_n), \dots, \alpha'_T(f_K, \xi_n)]^T;$$

$\rho(\Theta_n \cdot \xi_n, \mathbf{A}_3)$  – расстояние в метрическом пространстве комплексных КО ПЭМВ между экспериментально полученными  $\mathbf{A}_3$  и вычисленными теоретическими значениями  $\mathbf{A}_T$  КО поля ПЭМВ в области допустимых значений  $\xi_n \in \xi_{\text{доп}}$ ;  $q_n^{\dot{\epsilon}}, q_n^{\dot{\mu}}, q_n^t$  – веса, определяющие степень априорной информации о комплексных диэлектрических проницаемостях, комплексных магнитных проницаемостях и толщинах слоев МП;  $\dot{\epsilon}_n, \dot{\mu}_n, t_n$  – комплексные диэлектрические проницаемости, комплексные магнитные проницаемости и толщины слоев РППМ (КМ);  $\dot{\epsilon}_{n0}, \dot{\mu}_{n0}, t_{n0}$  – найденные с учётом априорной информации о составляющей среднее значение  $n$ -й составляющей вектора ЭФГП из области допустимых значений  $\dot{\epsilon}_{n0} \in [\dot{\epsilon}_{n\min}, \dots, \dot{\epsilon}_{n\max}]$ ,  $\dot{\mu}_{n0} \in [\dot{\mu}_{n\min}, \dots, \dot{\mu}_{n\max}]$ ,  $t_{n0} \in [t_{n\min}, \dots, t_{n\max}]$ ;  $\varpi$  – параметр регуляризации.

В соответствии с [10, 11] мнимые части  $\alpha''(f, \Phi_\alpha)$  экспериментальных комплексных КО учтены, как дополнительный параметр при минимизации целевой функции (6), путем представления их в виде параметрической функции:  $\alpha''(f, \Phi_\alpha) = \sum_{i=0}^n a_{n-i}^r f^i$ , где  $a_{n-i}$  – параметры, которые представляют собой коэффициенты регрессии данной функциональной зависимости мнимой части КО;  $f$  – текущая частота измерения;  $\Phi_\alpha = \{a_{n-i}\}$ ,  $i=0, \dots, n$  – вектор параметрических коэффициентов данной функциональной зависимости.

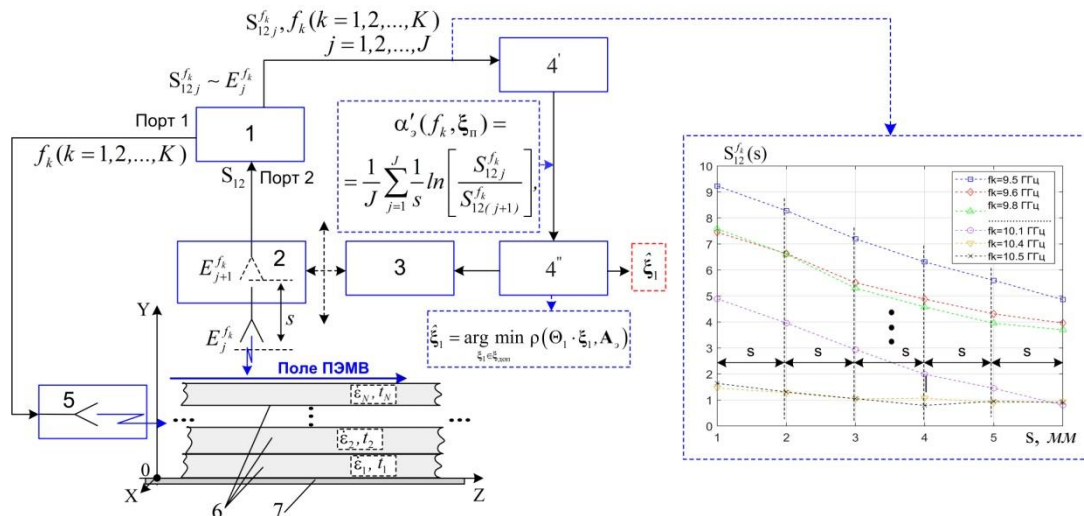
Фундаментальная особенность использования ПЭМВ заключается в том, что глобальный минимум целевой функции (6) довольно «острый» и расположенный в относительно узкой полосе оцениваемых ЭФГП. Выбором сред с конечным числом слоев и ограниченным диапазоном возможных значений их параметров выделяется компактный класс решений, обеспечивающий единственность решения обратной задачи при минимизации функционала (6). Регуляризация обратной задачи (6) сводится к выбору параметра регуляризации  $\varpi$  и весовых коэффициентов  $q_n^{\dot{\epsilon}}, q_n^{\dot{\mu}}, q_n^t$  [10, 11].

### Материалы и оборудование

Измерительный комплекс (ИК) для реализации метода определения ЭФГП многослойных материалов и покрытий реализован в виде стенда на основе векторного анализатора электрических цепей (ВАЦ) «Панорама Р4226» (АО «НПФ «Микран»»), системы перемещения приемной антенны (измерительного зонда) с числовым программным управлением и осциллографа KeysightDSO9404A (AgilentTechnologies) [11].

На рис. 2 представлена структурная схема и общая схема проведения эксперимента по оценке ЭФГП многослойных материалов и покрытий, где цифрами обозначено: 1 – векторный анализатор электрических цепей; 2 – приемная антенна с механизмом перемещения – 3; 4 –

электронная вычислительная машина (осциллограф Keysight DSO9404A); 5 – антенна возбуждения ПЭМВ; 6 – исследуемый образец многослойного покрытия; 7 – металлическая подложка; 8 – радиопоглощающий материал для «экранировки» приемной антенны (измерительного зонда); 9 – набор исследуемых образцов материалов; 10 – вариант отображения частотной зависимости коэффициентов передачи  $S_{12}$  на осциллографе.



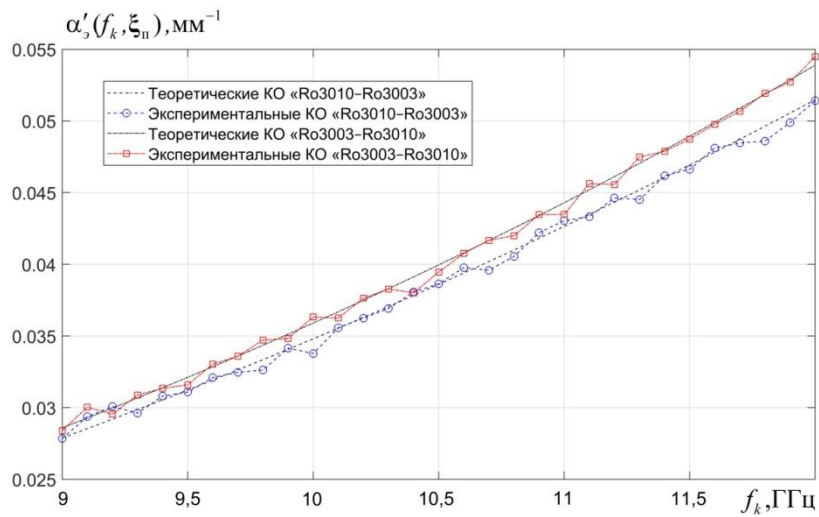
**Рис. 2. Структурная схема ИК, реализующего метод определения ЭФГП многослойных материалов и покрытий**

В качестве образцов материалов для формирования многослойных покрытий для натуральных экспериментов использовались промышленные СВЧ-ламинаты из одной серии на основе фторопласта FR-4 с керамическим наполнителем, предназначенные для изготовления СВЧ-печатных плат: RogersRo3010 ( $\epsilon = 10,2 - j0,022$ ,  $t = 0,64$  мм), RogersRo3006 ( $\epsilon = 6,15 - j0,012$ ,  $t = 1,28$  мм), RogersRo3003 ( $\epsilon = 3 - j0,003$ ,  $t = 0,13$  мм) [11, 15]. Данные материалы в частотном диапазоне 8–15 ГГц обладают практически частотно-независимыми значениями комплексной диэлектрической проницаемости [15]. Размеры образцов 200×100 мм.

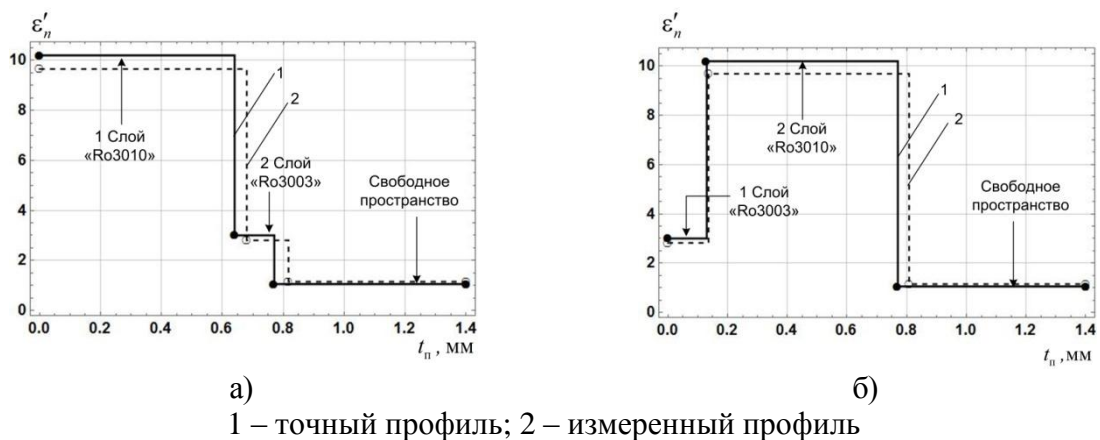
### Результаты и их обсуждение

Для исследования двухслойных покрытий изготовлено два образца покрытий на основе Rogers Ro3010 и RogersRo3003: «Ro3010–Ro3003» и «AD255C–Ro3010». Для оценки параметров двухслойных покрытий приняли полосу частот измерений  $\Delta f = 9 - 12$  ГГц с шагом по частоте 0,1 ГГц (количество частот  $K=31$ ). Параметры регуляризации для целевой функции (6) при определении параметров двухслойного покрытия следующие:  $\varpi = 0,0018$ ,  $q_1^{\epsilon'} = 1,8$ ,  $q_1^{\epsilon''} = 1,8$ ,  $q_1^t = 1,7$ ,  $q_2^{\epsilon'} = 1,8$ ,  $q_2^{\epsilon''} = 1,8$ ,  $q_2^t = 1,7$ . Перед началом проведения измерений проводили калибровку ИК на двухслойном покрытии из двух слоев стандартного образца ПММК равной толщины ( $t=0,5$  мм). Полученные экспериментальные частотные зависимости КО ПЭМВ покрытий «Ro3010–Ro3003» и «Ro3003–Ro3010» приведены на рис. 3.

Указанные частотные зависимости для данных покрытий различаются, что позволяет надежно определять порядок сочетания слоев. Далее по ним решалась обратная задача (6). На рис. 4а и 4б показан точный профиль действительной части комплексной диэлектрической проницаемости и полученный в результате решения обратной задачи (6) по экспериментальным данным с ИК.



**Рис. 3. Экспериментальные частотные зависимости КО поля ПЭМВ  $\alpha'_o(f_k, \xi_n)$  двухслойных покрытий на металлической подложке «Ro3010–Ro3003» и «Ro3003–Ro3010»**



**Рис. 4. Экспериментально полученные профили двухслойных покрытий на металлической подложке «Ro3010–Ro3003» (а) и «Ro3003–Ro3010» (б)**

Получены следующие экспериментальные относительные погрешности оценки комплексной диэлектрической проницаемости и толщин материалов Ro3010 и Ro3003 при их оценке в составе двухслойных покрытий с доверительной вероятностью 0,95:

– покрытие «Ro3010–Ro3003»:  $\hat{\epsilon}_{Ro3010} \approx 9,67 - j0,023$  ,  $\hat{\epsilon}_{Ro3003} \approx 2,84 - j0,0032$  (средние относительные погрешности оценки  $\delta_{\epsilon'} \approx 5,2\%$  ,  $\delta_{\epsilon''} \approx 6,1\%$  );  $\hat{t}_{Ro3010} \approx 0,68$  мм,  $t_{Ro3003} \approx 0,138$  мм (средняя относительная погрешность оценки толщин слоев 5,2 %);

– покрытие «Ro3003–Ro3010»:  $\hat{\epsilon}_{Ro3003} \approx 2,85 - j0,0031$  ,  $\hat{\epsilon}_{Ro3010} \approx 9,8 - j0,022$  , (средние относительные погрешности оценки  $\delta_{\epsilon'} \approx 5\%$  ,  $\delta_{\epsilon''} \approx 6\%$  );  $\hat{t}_{Ro3003} \approx 0,137$  мм,  $\hat{t}_{Ro3010} \approx 0,67$  мм (средняя относительная погрешность оценки толщин слоев 5 %).

Таким образом, получено устойчивое решение обратной задачи (6). Экспериментальные относительные погрешности оценки ЭФГП слоев соизмеримы между собой, даже с учетом наличия слоя с небольшой электрической толщиной.

При оценке ЭФГП двухслойных покрытий минимизация целевой функции (6) осуществлялась с помощью генетического алгоритма, реализованного в виде функции «ga» приложения GlobalSearch пакета Matlab.

### Выводы

Таким образом, на основе выявленных фундаментальных закономерностей, связывающих электрофизические и геометрические параметры плоскостойких многослойных МП с



характеристиками поля ПЭМВ, развит метод, обеспечивающий в рамках единого методического подхода проведение измерения параметров многослойных диэлектрических и магнетодиэлектрических МП в едином цикле измерения, в одних и тех же условиях (размещение на металлической подложке, конструкционном композиционном материале и др.).

Экспериментальные исследования двухслойных диэлектрических покрытий показали, что при ширине полосы частот измерений 9–13 ГГц обеспечиваются оценки действительных и мнимых частей слоев с относительной погрешностью не больше 6%, а толщин слоев не больше 5,5 % с доверительной вероятностью 0,95.

### Библиографический список

1. **Казанцева, Н.Е.** Перспективные материалы для поглотителей электромагнитных волн сверхвысококалостотного диапазона / Н.Е. Казанцева, Н.Г. Рывкина, И.А. Чмутин // Радиотехника и электроника. 2003. Т. 48. №2. С. 196–207.
2. **Lagarkov, A.N.** Dielectric properties of fiber-filled composites / A.N.Lagarkov, S.M.Matytsin, K.N.Rozanov, A.K.Sarychev // Journal of Applied Physics. 1998. vol. 84. № 7, pp. 3806–3814.
3. **Plamen I. Dankov** Experimental Characterization of Positive and Negative Dielectric Constants and Artificial Anisotropy of Meta-materials in the Microwave Range / Plamen I. Dankov // J. Phys.: Conf. Ser. 2020. V. 1598 (012002). DOI: 10.1088/1742-6596/1598/1/012002.
4. **Богданов, Ю.** Фольгированные диэлектрики – как выбрать оптимальный вариант для печатных плат ВЧ/СВЧ-диапазонов. Часть 1 / Ю.Богданов, В.Кочемасов, Е.Хасьянов // Печатный монтаж. 2013. № 2. С. 156–168.
5. **Михайлин, Ю.А.** Специальные полимерные композиционные материалы / Ю.А. Михайлин. СПб.: Научные основы и технологии, 2008. 660 с.
6. **Вендик, И.Б.** Метаматериалы и их применение в технике сверхвысоких частот (Обзор) / И.Б.Вендик, О.Г.Вендик // Журнал технической физики. 2013. Т. 83. вып. 1. С. 3–28.
7. **Слюсар, В.** Метаматериалы в антенной технике: основные принципы и результаты / В.Слюсар // Первая миля. 2010. №3–4. С. 44–60.
8. **Гринев, А.Ю.** Радары подповерхностного зондирования. Мониторинг и диагностика среди объектов // А.Ю.Гринев, В.С.Темченко, Д.В.Багно. М.: Радиотехника, 2013. 391 с.
9. **Kaz'min, A.I.** Evaluating the Accuracy of Reconstruction of the Electrical and Geometric Parameters of Multilayer Dielectric Coatings by a Multifrequency Radio-Wave Method for Slow Surface Electromagnetic Waves / A.I.Kaz'min, P.A.Fedyunin // Measurement Techniques, November 2020. Volume 63. No. 8. Pp 645–652. DOI: <https://doi.org/10.1007/s11018-020-01834-9>.
10. **Chernyshov, V.** Testing Electrophysical Parameters of Multilayer Dielectric and Magnetodielectric Coatings by the Method of Surface Electromagnetic Waves / V. Chernyshov, A. Kaz'min and P. Fedyunin // 2021 3rd International Conference on Control Systems, Mathematical Modeling, Automation and Energy Efficiency (SUMMA), 2021. Pp. 372–377, DOI: 10.1109/SUMMA53307.2021.9632211.
11. **Казьмин, А.И.** Методологические принципы определения электрофизических параметров материалов и покрытий со сложной внутренней структурой с помощью поверхностных электромагнитных волн / А.И. Казьмин // Дефектоскопия. 2022. № 3. С. 34–49.
12. **Уолтер, К.** Антенны бегущей волны. Пер. с англ., под общ. ред. А.Ф. Чаплина / К.Уолтер. М.: Энергия, 1970. 448 с.
13. **Ваганов, Р.Б.** Экспериментальное исследование структуры поверхностной электромагнитной волны в анизотропно проводящей ленте / Р.Б.Ваганов, И.П.Коршунов, Е.Н.Коршунова, А.Д.Олейников // Радиотехника и электроника. 2013. Т. 58. № 2. С. 136–142.
14. **Фелсен, Л.** Излучение и рассеяние волн / под ред. М.Л. Левина, пер. с англ. / Л.Фелсен, Н.Маркувиц. М.: Мир, 1978. Т.1. 546 с. Т.2. 550 с.
15. DataSheet RO3000® Series Circuit Materials RO3003™, RO3006™, RO3010™ and RO3035™ High Frequency Laminates / <https://www.rogerscorp.com/>.

## О МЕТОДАХ ПРОЕКТИРОВАНИЯ ФЕРРИТОВЫХ ЦИРКУЛЯТОРОВ НА СОСРЕДОТОЧЕННЫХ ЭЛЕМЕНТАХ

В.А. Козлов

ORCID: 0000-0002-9469-2284 e-mail: vkozlov@niiis.nnov.ru

Е.А. Михалицын

ORCID: 0000-0003-3144-3704 e-mail: mihalitsynea@gmail.com

А.В. Сорокин

ORCID: 0000-0002-7730-7129 e-mail: insav@rambler.ru

Филиал РФЯЦ-ВНИИЭФ «НИИИС им. Ю.Е. Седякова»

В докладе представлен анализ трех методов проектирования ферритовых циркуляторов на сосредоточенных элементах и выделены их преимущества, недостатки и область применения в комплексном процессе разработки ферритовых развязывающих устройств. Первый метод основан на феноменологической схмотехнической модели замещения ферритового Y-сочленения сосредоточенными индуктивностями собственных двухполюсников и является наиболее широко используемым в инженерной практике в настоящее время. Метод применим лишь к симметричным топологиям. Второй метод основан на идее схмотехнического замещения системы переплетенных проводников двенадцатью сосредоточенными связанными индуктивностями. Коэффициенты связи зависят от угла пересечения проводников, и такая система может быть использована для анализа асимметричных топологий Y-сочленения. В работе представлены результаты интеграции первых двух методов в схмотехнический САПР разработки СВЧ устройств. Третий подход – трехмерное электродинамическое моделирование в САПР СВЧ устройств, - базируется на первых двух методах и позволяет более детально проанализировать влияние паразитных параметров топологии и провести её оптимизацию. Приводится анализ результатов применения трех методов к проектированию оптимизированной топологии циркулятора дециметрового диапазона длин волн.

**Ключевые слова:** циркулятор на сосредоточенных элементах, система переплетённых проводников, индуктивность связи скрещивающихся проводников, трехмерное электродинамическое моделирование.

### THE FERRITE LUMPED ELEMENTS CIRCULATORS DESIGN METHODS

V.A. Kozlov, E.A. Mihalitsyn, A.V. Sorokin

Branch of federal state unitary enterprise «Russian federal nuclear center - All-russian scientific-research institute of experimental physics» «Scientific-research institute of measuring systems  
n.a. Yu.E. Sedakov»

**Purpose:** The authors provides an analysis and application results of three methods of design L-band lumped element ferrite circulators.

**Design / methodology / approach:** There are three design procedures to be considered here. The first is phenomenological schematic design based on ferrite Y-junction substitution on the lumped eigen-inductivities. The second design method is based on mutual coupling calculation of cross stripes and schematic analysis of twelve ideal transformers representing ferrite Y-junction. 3D Electromagnetic modeling of ferrite Y-junction in 3D CAE of high frequency devices is the third approach which allows comprehensive analysis of parasitic effects and topology optimization.

**Findings:** The paper finds out optimal ferrite Y-junction topology which obtains high quality performance of designed L-band circulator.

**Research limitations/implications:** The first approach is limited by symmetric lumped element circulators and widely used in practical engineering. The second approach could be used in design procedures of asymmetric Y-junction circulators. Methods are integrated in CAE of high frequency devices.

**Originality/value:** The value of the whole paper is the complex lumped element circulators design procedure based on three approaches presented in the paper.

**Key words:** lumped element circulator, crossover network of traces, mutual inductance of cross stripes, 3D-electomagnetic modeling.

## Введение

Ферритовые развязывающие устройства (ФРУ) уже давно применяются в радиотехнических системах (РТС) различного назначения 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. Несмотря на быстрое развитие полупроводниковых компонентов, которые по своим функциональным возможностям во многих случаях успешно конкурируют с ФРУ, их применение неизменно остается актуальным в условиях радиационных воздействий и высоких уровней мощности.

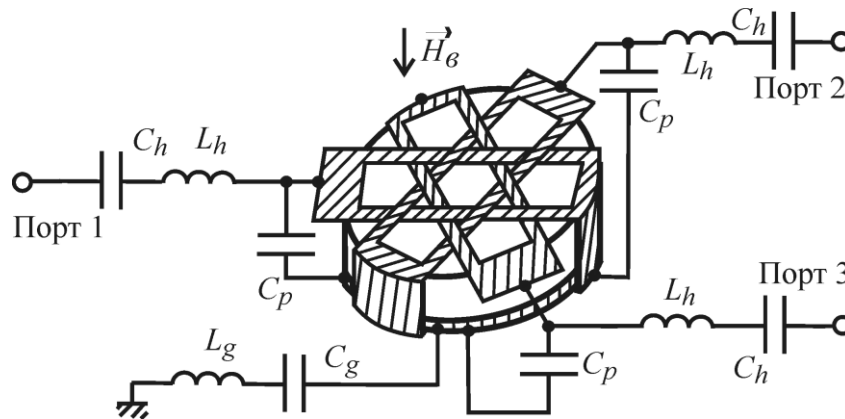
Наименьшими габаритами в метровом и сантиметровом диапазонах длин волн характеризуются циркуляторы на сосредоточенных элементах. Их отличие от других типов ферритовых циркуляторов [1], которое обеспечивает их миниатюрность, заключается в структуре Y-соединения, где используется система переплетенных проводников, расположенная между двумя ферритовыми элементами.

При проектировании симметричных циркуляторов на сосредоточенных элементах традиционно используют подход, основанный на эквивалентной схеме замещения ферритового Y-соединения системой связанных сосредоточенных индуктивностей [2]. Для конструкций с выраженной асимметрией переплетения, когда проводники пересекаются под углами отличными от 120 градусов, такой метод не применим. Для анализа таких схем в данной работе рассмотрен подход, основанный на идее схмотехнического замещения системы переплетенных проводников двенадцатью сосредоточенными связанными индуктивностями. Коэффициенты связи зависят от угла пересечения проводников, и поэтому такая система может быть использована для анализа асимметричных топологий Y-соединения. Неточности проектирования первым и вторым методами, связанные с наличием паразитных индуктивностей и емкостей связи могут быть устранены в рамках третьего подхода, основанного на трехмерном моделировании в САПР СВЧ устройств. В данной работе описаны ключевые моменты процесса проектирования узкополосного циркулятора дециметрового диапазона длин волн.

### 1. Феноменологический метод расчета ферритовых циркуляторов

Типовая эквивалентная схема циркулятора на сосредоточенных элементах приведена на рис.1. Проводники в точках пересечения изолированы друг от друга диэлектрическими прокладками и подключены к заземлению ферритовых дисков. В каждом плече Y-соединения установлены шунтирующие настроечные емкости  $C_p$ . Внешнее статическое магнитное поле  $H_e$  направлено перпендикулярно плоскости ферритов. В схему полосно-расширяющие согласующие  $L_n C_n$  цепи и цепь общего контура  $L_g C_g$ .

При феноменологическом проектировании этого типа циркуляторов принимаются базовые физические допущения, предложенные в [3, 4]. Так, например, предполагается, что в случае возбуждения только одного порта высокочастотное магнитное поле однородно распределено по поверхности ферритовых дисков и направлено перпендикулярно проводникам, а также что вариация интенсивности высокочастотного электрического поля по длине проводников отсутствует, а энергия электрического поля много меньше в сравнении с магнитной. При этом для типов возбуждения с правым и левым направлениями вращения магнитного поля тригонально симметричная система переплетения проводников обеспечивает формирование электромагнитного поля с круговой поляризацией во всем объеме ферритовых дисков с заземлением.



**Рис.1. Ферритовое Y-сочленение с системой переплетенных проводников (верхний феррит не показан)**

Для анализа симметричных схем циркуляторов на сосредоточенных элементах применяется метод собственных возбуждений [1, Chapter 12], согласно которому рассматриваемая тригонально-симметричная структура заменяется тремя собственными двухполюсниками, соответствующими синфазному собственному вектору и векторам собственных возбуждений с циркулярными: правым и левым, направлениями вращения плоскости поляризации напряженности магнитного поля в ферритах. Идеальные условия циркуляции на частоте циркуляции  $f_c$  заключаются в симметричном под углами 120 градусов по отношению друг к другу взаимном расположении коэффициентов отражения собственных двухполюсников на комплексной плоскости.

Для узкополосного циркулятора, настроенного на частоту  $f_c$  в отсутствие полосно-расширяющих цепей процедура расчета схемных элементов может быть представлена в простом аналитическом виде. Пусть помимо частоты будут заданы, как это бывает на практике, волновое сопротивление подводящих линий  $\rho_0$ , намагниченность насыщения феррита  $M_s$  и геометрия ферритового элемента с эквивалентной индуктивностью  $L_0$ . Тогда из общей методики [5] можно вывести, что напряженность внутреннего магнитного поля  $H_0$  может быть рассчитана из выражения:

$$\sigma = \sqrt{p\rho_0 / (\sqrt{3}\pi f_c L_0) + 1} - p, \quad (1)$$

а шунтирующая настроечная емкость определится по формуле:

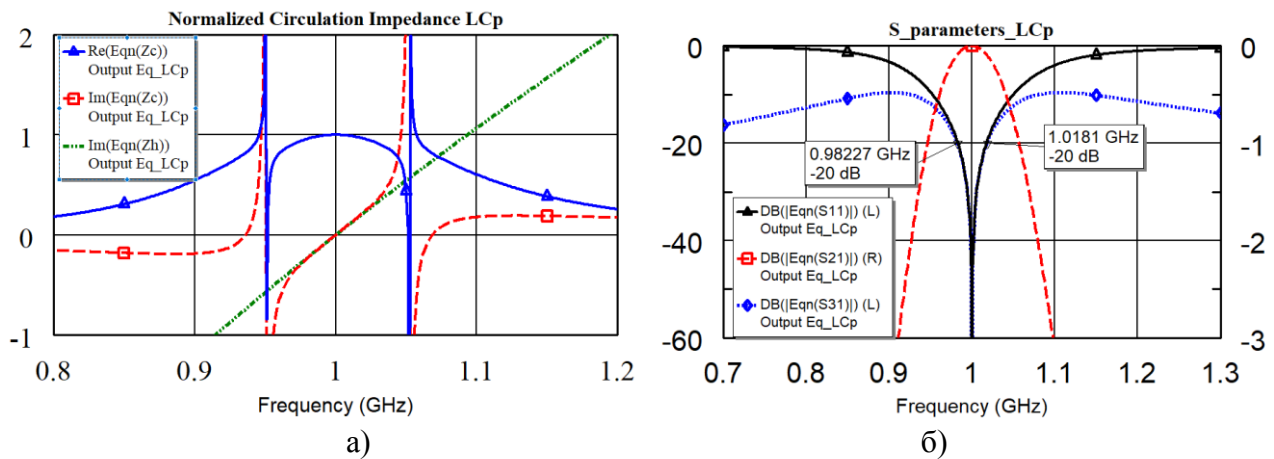
$$C_p = \frac{2\rho_0 - \sqrt{3(2\pi f_c L_0)^2 + 4\sqrt{3}\pi\rho_0 f_c L_0}}{3\rho_0 2\pi f_c^2 L_0}, \quad (2)$$

где  $\sigma = \gamma H_0 / f$  - относительное поле подмагничивания,  $p = \gamma 4\pi M_s / f$  - параметр намагниченности,  $\gamma = 2.8$  МГц/Э - гиромангнитное отношение,  $H_g = H_0 + N 4\pi M_s$ ,  $N$  - фактор размагничивания по нормали к поверхности феррита  $N \approx 0.8 \dots 0.95$  [1].

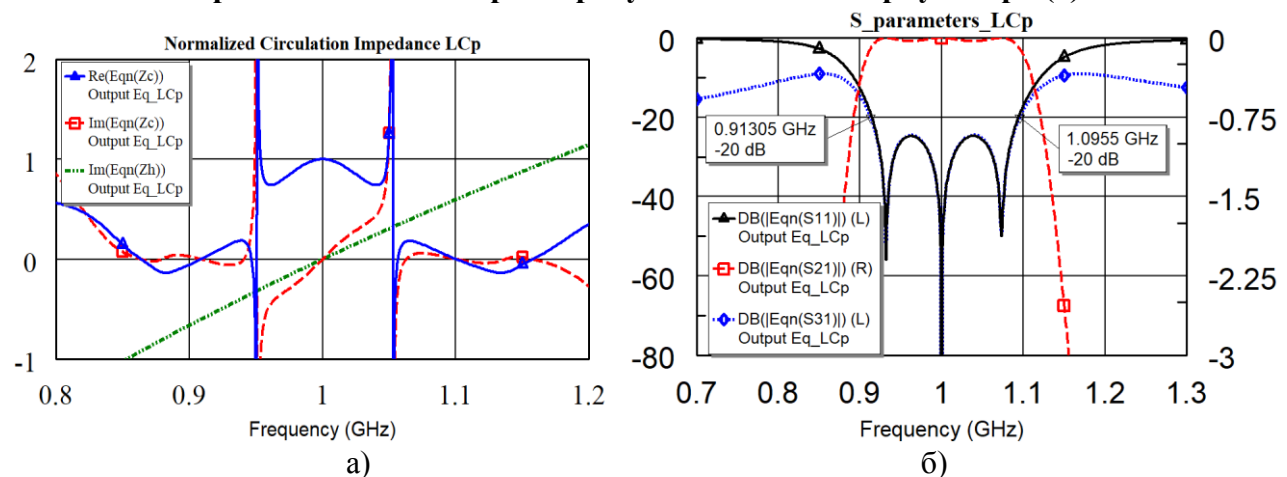
Методика автоматизированного проектирования широкополосных циркуляторов, основанная на феноменологическом подходе и методе анализа импеданса циркуляции описана в [6]. С идейной точки зрения, в схему добавляются последовательные согласующие LC цепи в каждое плечо циркулятора и цепь общего контура так, чтобы центральная частота равнялась частоте циркуляции  $f_c$ , а средний наклон мнимой части импеданса циркуляции Y-циркулятора в как можно большей полосе частот был максимально близким к наклону мнимой части импеданса согласующей цепи (рис.2а, рис.3а). Импеданс циркуляции - это комплексная нагрузка каждого из портов симметричного циркулятора, при которой имеет место идеальная циркуляция. Численное моделирование показало, что характеристические

сопротивления LC цепи общего контура  $\rho_g$  и согласующих цепей  $\rho_h$  должны удовлетворять приближенному соотношению:  $\rho_h \approx 3\rho_g$ . Тогда действительная часть нормированного импеданса циркуляции устройства будет близка к единице, мнимая часть – к нулю. При проектировании удобно использовать схемотехнический САПР СВЧ устройств.

Частотные характеристики узкополосного и широкополосного циркуляторов со следующими исходными проектными параметрами:  $M_s=1700$  Гс,  $f_c=1$  ГГц,  $L_0=0.47$  нГн, и рассчитанными по формулам (1) и (2) настроечными параметрами  $H_0=1750\Omega$ ,  $C=17.9$  пФ,  $\rho_g = 1.24\rho_0$ ,  $\rho_h = 3.13\rho_0$ , представлены на рис. 2 и рис. 3. Характеристики широкополосного устройства имеют трехгорбый вид, тогда как для узкополосного циркулятора имеет место лишь одна частотная точка выполнения условий идеальной циркуляции. Как следует из представленных графических зависимостей S-параметров, при фиксированной величине  $L_0$  теоретически реализуемая полоса рабочих частот по уровню обратных потерь 20 дБ составляет 18%, тогда как для узкополосного устройства достижима лишь полоса рабочих частот не более чем в 3.6 %.



**Рис.2. Частотная зависимость действительной и мнимой частей нормированного импеданса циркуляции и мнимой части импеданса нагрузки (а) и соответствующие рассчитанные S-параметры узкополосного циркулятора (б)**



**Рис.3. Частотная зависимость действительной и мнимой частей нормированного импеданса циркуляции и мнимой части импеданса нагрузки (а) и соответствующие рассчитанные S-параметры широкополосного циркулятора (б)**

Приведенные на рис.3а графики частотной зависимости действительной и мнимой частей нормированного импеданса циркуляции для широкополосной схемы циркулятора без согласующей цепи в сравнении с графиком частотной зависимости мнимой части импеданса

последовательной  $L_n C_n$  цепи, показывают близость наклона соответствующих кривых: мнимых частей импедансов друг к другу, а действительной части нормированного импеданса циркуляции – к единице.

## 2. Схемная модель системы переплетенных проводников на локальных связях

Второй подход к проектированию циркуляторов на сосредоточенных элементах основан на идее схемотехнического замещения системы переплетенных проводников двенадцатью сосредоточенными связанными индуктивностями и емкостями связи (рис.4). Взаимный коэффициент связи двух скрещивающихся полосковых линий в ферритовой среде зависят от угла пересечения проводников [5]:

$$M \approx ik \frac{1 - \exp\{-\beta w(k/\mu + ictg \theta)\}}{1 - \exp\{-k/\mu\beta w\}} \begin{cases} y_2(1 - y_1/2H), y_1 \geq y_2 \\ y_1(1 - y_2/2H), y_2 \geq y_1 \end{cases} \quad (5)$$

где  $w_{1e} = w_{2e} = w$  - эффективная ширина полосковых линий,  $\beta = k_0 \sqrt{\epsilon \mu}$ ,  $\theta$  - угол между скрещивающимися проводниками,  $y_1, y_2$  - расстояния от заземляющей поверхности до каждого из скрещивающихся проводников,  $2H$  - толщина ферритовой сборки,  $k, \mu$  - параметры тензора магнитной проницаемости феррита.

Схема замещения строится на основе концепции локальной связи, определяемой как взаимную связь между конечными участками перекрещивающихся полосковых линий. Эквивалентная схема рис. 4 с дополнением формулой (5) обеспечивает альтернативный механизм описания невзаимности циркулятора, который можно описать следующим образом. Пара полосков, выходящих из первого порта пересекается с полосками выходящими из двух других портов. Однако взаимная индуктивность скрещивающихся полосков слева и справа от порта различаются знаком, что приводит к различному характеру невзаимности возбуждаемых мод, соответственно, с левым и правым направлениями вращения плоскости поляризации возбуждаемых волн вблизи цепочки полосковых перекрестий. Хотя эти моды и напоминают распределенные вращающиеся резонансные колебания в ферритовом диске, но они имеют фундаментально другую физическую природу – волны возбуждаются цепочкой локализованных невзаимных связей на перекрестиях проводников. При этом взаимная индуктивность существенно зависит от угла между пересекающимися проводниками, в то время как взаимная емкость, главным образом, определяется площадью перекрытия проводников. Вариация угла пересечения полосков позволяет менять проектную величину индуктивности  $L_0$  для каждого из портов циркулятора, влияя на рабочую полосу частот.

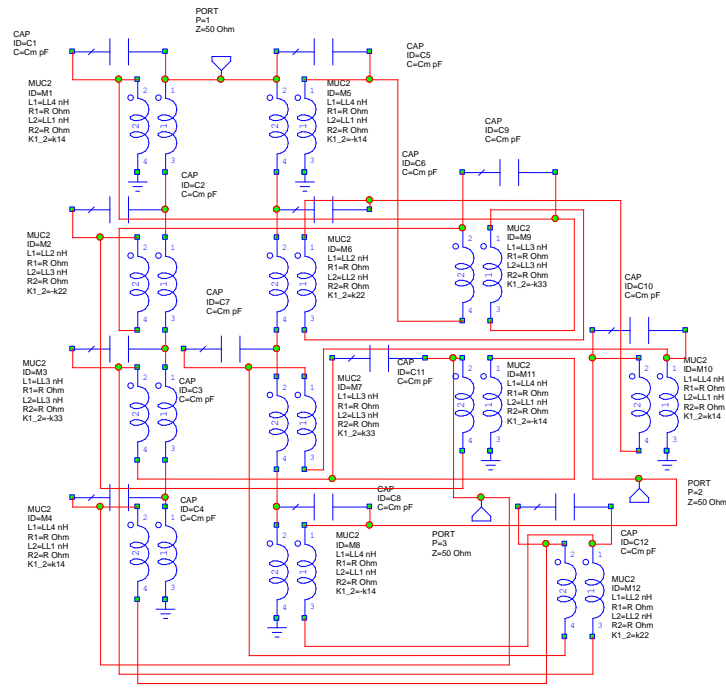


Рис. 4. Схемная модель системы переплетенных проводников

### 3. Методика экстракции параметра индуктивности системы переплетенных проводников из S-параметров

Реализация проектной величины индуктивности  $L_0$  системы переплетенных проводников ферритового Y-разветвления является наиболее технологически сложной и трудоемкой задачей проектирования циркуляторов на сосредоточенных элементах. Трехмерное электродинамическое моделирование топологии ферритового Y-разветвления (рис.5) позволяет учесть влияние всех геометрических параметров конструкции на рабочие характеристики устройства. Для обеспечения связи геометрических параметров со проектными параметрами целесообразно развитие методик экстракции номиналов элементов схемы замещения ферритового Y-циркулятора из результатов моделирования.

Может быть предложена следующая методика экстракции индуктивности  $L_0$  из результатов расчета S-параметров Y-разветвления с системой переплетенных проводников. Проводится моделирование или измерение S-матрицы рассеяния трехпортового устройства рис.5 в САПР 3D моделирования. Нормированная матрица импеданса ферритового Y-разветвления  $\mathbf{N}$  при большом поле подмагничивания задается в виде [6]:

$$\mathbf{N} = jX \begin{bmatrix} 1 & -1/2 & -1/2 \\ -1/2 & 1 & -1/2 \\ -1/2 & -1/2 & 1 \end{bmatrix}, \quad (6)$$

где  $X = L_0 2\pi f [1 + p/\sigma] / \rho_0$ .

Имея в виду, что помимо индуктивной составляющей импеданса циркулятора имеет место паразитная емкостная, обозначим действительную часть нормированной проводимости паразитной шунтирующей емкости  $C'$  в форме  $Y = 2\pi f C' \rho_0$ . Тогда матрица импеданса циркулятора с шунтирующей емкостью в каждом порту циркулятора определится в матричной форме:

$$\mathbf{Z} = (j\mathbf{Y}\mathbf{N} + \mathbf{E})^{-1}\mathbf{N}, \quad (7)$$

где  $\mathbf{E}$  - единичная матрица.

Алгебраические вычисления приводят к следующему выражению для элемента  $Z_{11}$  итоговой матрицы импеданса

$$Z_{11} = -\frac{2jX}{3XY - 2}. \quad (8)$$

Не трудно видеть, что при отсутствии паразитной шунтирующей емкости  $Z_{11} = X$ . Моделирование S-матрицы циркулятора в более широкой полосе частот позволяет найти резонансную частоту  $f_r$ , при которой  $Z_{11}$  обращается в бесконечность. Тогда паразитная шунтирующая емкость на землю  $C_p$  может быть выражена в виде:

$$C_p = \left[ 3/2L_0\omega_r^2(1 + p_r\sigma_r/(\sigma_r^2 - 1)) \right]^{-1}, \quad (9)$$

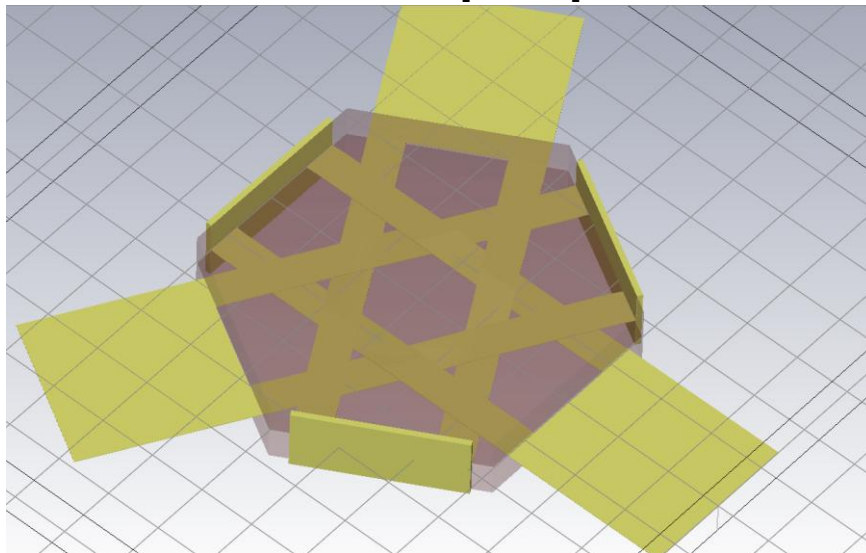
а параметр  $X$  определится по формуле:

$$X = \left[ 1 - (f/f_r)^2 \Omega \right] \text{Im}(Z_{11}), \quad (10)$$

где  $\Omega \approx \frac{1 + p/\sigma}{1 + p_r/\sigma_r}$ , а  $\sigma_r = \gamma H_0 / f_r$  и  $p_r = \gamma 4\pi M_s / f_r$ .

В итоге искомый параметр индуктивности рассчитывается по формуле:

$$L_0 = \frac{\rho_0 X}{2\pi f [1 + p/\sigma]}. \quad (11)$$



**Рис. 5. 3D-модель ферритового Y-разветвления с подводящими линиями**

На рис. 6а представлены результаты экстракции параметра  $L_0$  из топологии Y-разветвления рис.5:  $H_0=20000$ Э,  $f_r \approx 3.3$ ГГц. Оценка величины паразитной емкости согласно (9) по результатам экстракции дает результат:  $C_p \approx 3$  пФ. На рис. 6б приводятся результаты расчета S-параметров узкополосного циркулятора, построенного на основе данной топологии. Значения настроечных параметров:  $H_0=1750$  Э,  $C=11.3$ пФ,  $\rho_0=133$  Ом, полоса рабочих частот - 2%. Частотные характеристики заметно хуже полученных с применением феноменологического подхода, что объясняется наличием паразитных конструктивных емкостей и индуктивностей. Наличие паразитной последовательной индуктивности для системы переплетенных проводников и нестрогое соответствие комбинации величин  $H_0$ ,  $L_0$ ,  $C$ , а также наличие частотной зависимости собственной индуктивности (рис.6а) привели к необходимости подстройки также и волнового сопротивления портов циркулятора. Таким образом, процедура настройки циркулятора на практике с необходимостью требует наличия четвертьволновых подстроечных элементов согласования.



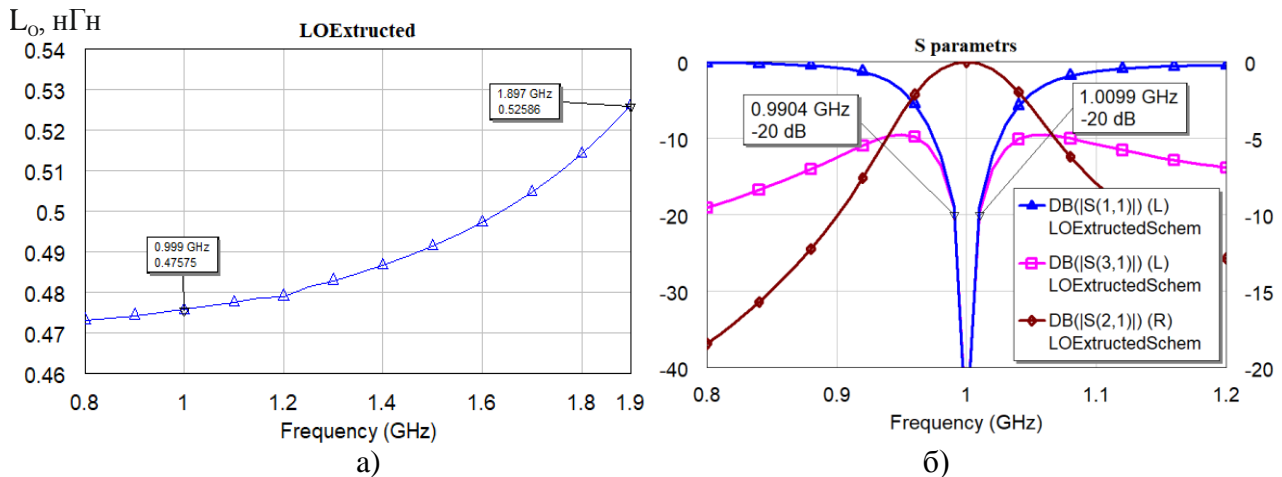


Рис. 6. Результат экстракции  $L_0$  из результатов моделирования (а), частотные характеристики соответствующего узкополосного циркулятора (б)

### Выводы

В представленной работе описаны основные методы проектирования ферритовых циркуляторов на сосредоточенных элементах. Приведены основные соотношения для феноменологического проектирования узкополосных и широкополосных циркуляторов, а также базовая идеологическая модель анализа асимметричных топологий системы на переплетенных проводниках. Подробно представлена новая методика экстракции проектного параметра циркулятора  $L_0$  из результатов трехмерного электродинамического моделирования S-параметров системы на переплетенных проводниках. Было показано хорошее соответствие результатов расчета S-параметров узкополосного циркулятора, данным, полученным по итогам трехмерного моделирования в САПР.

### Библиографический список

1. **Helzajn, J.** The Stripline Circulator Theory and Practice. John Wiley & Sons, 2008, 604 p.
2. **Сорокин, А.В.** Проектирование малогабаритных ферритовых развязывающих устройств на сосредоточенных элементах // Диссертация на соискание ученой степени кандидата технических наук. Нижний Новгород, 2014, 293 с.
3. **Konishi, Y.** Lumped element circulators // Transactions on Magnetics, 1975, vol. 11, no. 5, pp. 1262-1266.
4. **Konishi, Y.** Lumped element Y circulator. Transactions on Microwave Theory and Techniques, 1965, vol. 13, no. 6, pp. 852-864.
5. **Бородин, В.Н., Козлов, В.А., Михалицын, Е.А., Сорокин, А.В.** Автоматизированное проектирование ферритовых циркуляторов на сосредоточенных элементах. Физика волновых процессов и радиотехнические системы. 2020. Т.23, N4, с.74-84
6. **Dong, H.** Broad-band and High-isolation performance of a Lumped-element circulator: theory and experiment // A Dissertation Presented in Partial Fulfillment of the Requirements for the Degree of Doctor of Philosophy. 2014. 118 p.

## ВНЕДРЕНИЕ СЕЛЕКТИВНОЙ ПАЙКИ ПРИ СЕРИЙНОМ ПРОИЗВОДСТВЕ ЭЛЕКТРОННЫХ МОДУЛЕЙ

Г.П. Суровикина

ORCID: 0000-0001-8634-7053 e-mail: rustblaster@bk.ru

И.О. Усова

ORCID: 0000-0002-8788-4093 e-mail: usova\_i@mail.ru

С.В. Кузнецова

ORCID: 0000-0002-7326-5128 e-mail: svlika@mail.ru

Филиал РФЯЦ-ВНИИЭФ «НИИИС им. Ю.Е. Седатова»

В докладе представлен оптимизированный технологический маршрут сборки высокоплотных печатных узлов суперкомпьютерных блоков. Оптимизация осуществлена за счет разработки процесса автоматизированной селективной пайки многорядных разъемов и соединителей, полностью заменяющего традиционные процессы ручного монтажа с помощью паяльника. В работе отражена методика подбора оптимального термопрофиля позволяющая получить полное протекание припоя в отверстия плат разной конструкции.

**Ключевые слова:** высокоплотный монтаж, селективная пайка, изделия монтируемые в отверстия, пайки соединителей и разъемов

## INTRODUCTION OF SELECTIVE SOLDERING IN SERIAL PRODUCTION OF ELECTRONIC MODULES.

G.P. Surovikina, I.O. Usova, S.B. Kuznetsova

Branch of federal state unitary enterprise «Russian federal nuclear center - All-russian scientific-research institute of experimental physics» «Scientific-research institute of measuring systems n.a. Yu.E. Sedakov»

**Purpose:** Optimization of the board assembly route by automating the operation of mounting connectors into holes using selective soldering. Development of the optimal thermal profile for high-quality flow of solder into the holes.

**Design/methodology/approach:** A method is proposed for determining the soldering modes of ERI-connectors mounted in holes in an automated way, instead of manual installation with a soldering iron with further control of solder joints.

**Findings:** In the work, an analysis of the method of automated and manual soldering was carried out, the design was worked out. Aspects of the relevance of automation of soldering into holes and its further introduction into production are described.

**Research limitations/implications:** Continuation of work aimed at automating assembly processes.

**Originality/value:** The method of processing the selective soldering thermal profile allows to guarantee the receipt of stable quality of solder joints, connectors, connectors and other components of printed circuit boards mounted in holes of any design.

**Keywords:** high-density mounting, selective soldering of products mounted in holes, soldering of connectors and connectors

Технология монтажа в отверстия представляет собой метод монтажа электронных компонентов на печатную плату (далее – ПП), при котором выводы компонентов устанавливаются в сквозные отверстия платы и припаиваются к монтажному пояску и металлизированной внутренней поверхности отверстия. Эта технология является родоначальником подавляющего большинства современных технологических процессов сборки электронных модулей. Несмотря на распространение технологии поверхностного монтажа, есть области электроники, где технология монтажа в отверстия по сей день является доминирующей технологией сборки: силовые устройства, блоки питания, высоковольтные схемы мониторов и других устройств в областях авионики, автоматики АЭС и т.п. [1]. В основном изделия, монтируемые в отверстия (далее ИМО) – это электромеханические реле, разъемы, большие переменные резисторы, соединители разных типов, контакты, операционные усилители.

Стандартный технологический процесс монтажа ИМО состоит из следующих типовых этапов:

- подготовка выводов ИМО (если по требованиям конструкторской документации (далее КД) требуются формовка и/или обрезка);
- лужение выводов требуемым припоем (по КД);
- установка компонентов в отверстия посадочных мест на плате;
- пайка;
- отмывка;
- контроль качества паяных соединений.

В результате данных действий формируется паяное соединение, удовлетворяющее следующим требованиям [2,3]:

- форма паяных соединений должна быть скелетной, с вогнутым галтелями (что соответствует краевому углу смачивания менее  $90^\circ$ ) по границе припоя и без его избытка;
- она должна позволять визуально просматривать через тонкие слои припоя контуры входящих в состав соединения монтажных элементов.

Паяные соединения признаются не удовлетворительными, если в ходе визуально-оптического осмотра (при увеличении  $4-8\times$ ) выявляется следующее:

- несмачиваемость (угол смачивания паяемых поверхностей более  $90^\circ$ );
- неполное заполнение припоем металлизированного отверстия (припой не виден в отверстии);
- перемыкания выводов ИМО;
- нарушение целостности корпуса ИМО;
- разрушение вывода ИМО;
- отслоение монтажного пояска отверстия, расслоение ПП;
- наличие паяльных остатков;
- соединения темно-серого цвета с дендритной шероховатой поверхностью и крупными игольчатыми образованиями.

Наиболее популярный и широко распространенный технологический вариант реализации монтажа ИМО – монтаж с помощью паяльника. Много лет монтажники впаивали ИМО в отверстия двусторонних печатных плат с паяемым покрытием ГорПОС61. Операция выверена годами, набор припоев известен, режимы пайки фактически не изменялись.

В последние 20 лет конструкции электронных модулей претерпели существенные изменения:

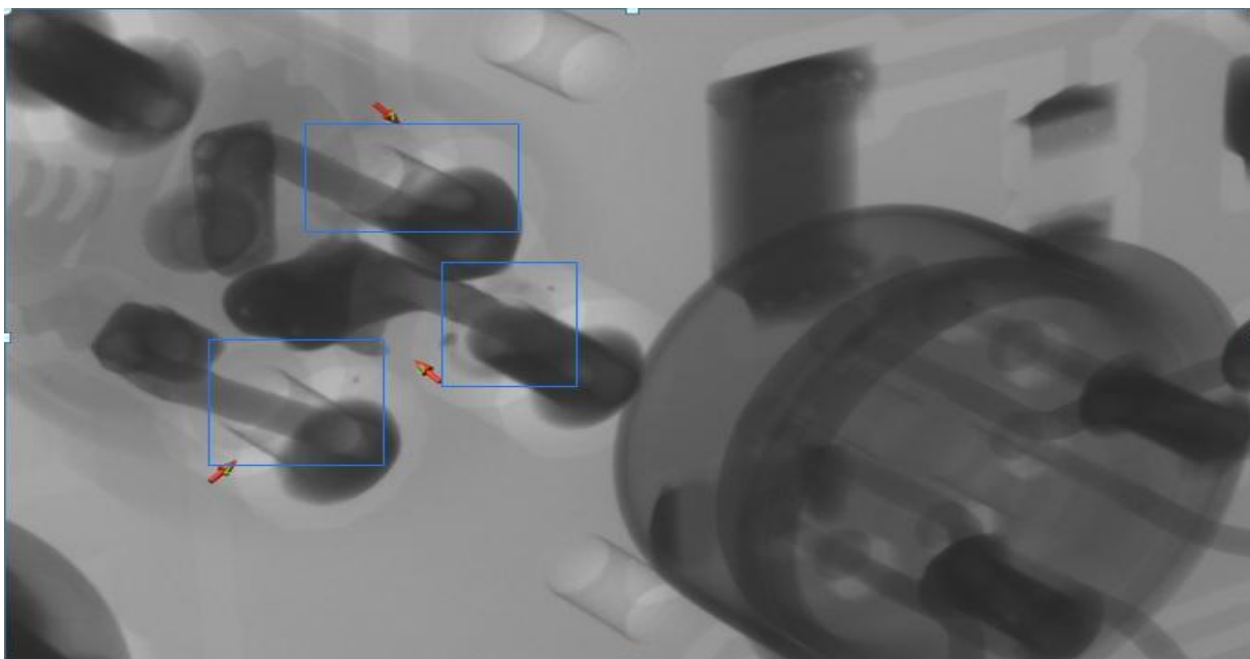
- появились многослойные печатные платы (до 28 слоев), платы с теплоотводящими слоями, а также платы, топологический рисунок слоев которых содержит металлизированные полигоны большой площади;
- схемотехнические решения часто реализуются таким образом, что в матрице отверстий посадочного места под один компонент часть отверстий во внутренних слоях платы могут быть соединены с разным количеством проводников на разных слоях платы.

Обеспечить полное протекание припоя в таких отверстиях, не меняя традиционной технологии пайки, невозможно. При этом высок риск получения некачественного паяного соединения, так как монтажник не имеет информации о внутреннем строении платы в области конкретного отверстия;

- наряду с традиционным оловянно-свинцовым покрытием, на платах и компонентах появились новые паяемые покрытия (с 2006 года, в связи с введением декларации RoHS[4]) – иммерсионное золото, иммерсионное олово, иммерсионное серебро, бессвинцовый сплав SAC305.

- увеличилось количество выводов разъёмов, соединителей, появились микросхемы в корпусах PGA с количеством штыревых выводов свыше 1000 штук, уменьшился шаг между штыревыми выводами.

Для пайки таких плат и компонентов, кроме паяльника, в типовом маршруте появились новые операции (предварительная сушка многослойных ПП) и дополнительное оборудование (подогревной столик, обеспечивающий дополнительный общий подогрев платы во время монтажа ИМО, чтобы обеспечить протекание припоя в отверстия плат сложной структуры, а также приспособления – стойки, чтобы не допускать касания корпусов компонентов разогретой поверхности столика). При изготовлении приборов ответственного назначения в случаях, когда невозможно выполнить визуально-оптический контроль протекания припоя (корпуса ИМО установлены вплотную к поверхности платы), успешно используются средства рентгенодефектоскопии.



**Рис.1. Иллюстративный снимок паяных соединений подстроечного резистора, полученный в ходе рентгенодефектоскопического контроля в режиме углового сканирования. Наблюдается отсутствие полного протекания припоя в отверстиях**

Ряд температурно-временных параметров режимов пайки даже для припоя одной марки значительно расширился. От монтажника требуется больше знаний, навыков и степени ответственности, чем 20 лет назад. Соответственно, на порядок возросло и влияние человеческого фактора на правильность выбора и соблюдения режимов пайки, особенно в случае серийного производства изделий.

Существует более надежный и стабильный вариант монтажа ИМО – это селективная пайка. Селективная пайка – метод пайки, при котором воздействию припоя подвергаются только места определенных будущих паяных соединений (как в случае с паяльником), при

этом пайка каждого соединения происходит по заданной программе. Операции подогрева, флюсования и пайки здесь реализуются в одной установке:

А) Нагрев платы сверху и снизу обеспечивается шестью нагревателями, оснащёнными динамическим кассетным излучателем. Ширина поля нагрева регулируется при помощи включения и выключения нагревателей, работающих независимо друг от друга. А значит, есть возможность избежать перегрева чувствительных к температуре компонентов.

Б) Флюсованиереализуетсямелкодисперсным распылением в автоматическом режиме с помощью микрокапельногофлюсователя. Такое распыление позволяет качественно локализовать нанесение флюса на каждое паяемое отверстие с точной дозировкой флюса, даёт возможность существенно уменьшить содержание «флюсового тумана» и уменьшить попутное загрязнение соседних компонентов.

В) В роли жала паяльника здесь выступает локальная мини-волна, формируемая отдельным тигелем определенного диаметра. Стабильность высоты волны в процессе пайки обеспечивается через подачу расплавленного припоя на головку посредством электромагнитного насоса.

Селективная пайкакак отдельная сборочно-монтажная технология не упоминается ни в отечественных технических условиях (далее – ТУ) или научно-технической документации (далее – НТД), ни в технических описаниях на импортные компоненты (datasheets). При отработке режимов селективной пайки исходили из тех же температурно-временных ограничений, которые используются при отработке режима пайки ИМО традиционным паяльником.Режим пайки конкретного ИМО определяется:

- рекомендациями ТУ (или datasheets, в зависимости от производителя изделия) на данный компонент (температурно-временными характеристиками, рекомендациями по применяемым материалам, допустимым типам воздействия, паяемым покрытиям выводов);

- требованиями КД;

- особенностями конструкции платы (составом паяемого покрытия, толщиной, структурой и материалом слоев, количеством и площадью проводников, соединенных с монтажным отверстием и т.п.);

- требованиями, изложенными в соответствующей НТД.

Освоение оборудования и внедрение технологии селективной пайки потребовалось для оптимизации маршрута сборки серийных изделий – плат адаптерного и коммутаторного блоков. Это современные конкурентоспособные электронные модули (далее ЭМ) с высокой плотностью монтажа и широким спектром устанавливаемых компонентов. Основу ЭМ составляют многослойные (18 и 28 слоев соответственно) печатные платы толщиной 1,7 и 2,8 мм из ламината типа FR5 с паяемым покрытием иммерсионное золото. На платы, кроме ПМИ широкого спектра, устанавливаются соединители импортного производства в бессвинцовом исполнении с паяемым покрытием иммерсионное олово. Конструкция плат была адаптирована под автоматизированные сборочные процессы: были рассчитаны и выполнены технологические поля шириной 10 мм, топологический рисунок верхних слоев плат содержал по 3 реперных знака в виде 2 кругов диаметром 1 мм и одного квадрата со стороной 1 мм, размещенных по диагонали платы в максимальной близости от ее краев. Высота корпусов компонентов над ПП не превышала 20 мм.

При отработке режима пайки ИМО на данные платы (как ручной, так и автоматизированной) исходили из следующих параметрических ограничений:

- температура пайки – не выше указанной в ТУ и в НТД (для самого распространенного припоя ПОС 61 для электрорадиоизделий (далее – ЭРИ) типа микросхем, реле, конденсаторов – не более 260°C, для разъемов и соединителей – не выше 280 – 290°C) [5]. Для импортных соединителей в бессвинцовом исполнении – не выше 300 °C. С учетом вышесказанного, а также зарубежного происхождения ИМО, температура пайки определялась – для мини-волны 285 °C, для паяльника – (280±10) °C;

- время пайки – не более 3 – 5 с для отечественных ИМО, и не более 10 с для всех соединителей и импортных ИМО;

- температура дополнительного нагрева многослойной платы при пайке – не более  $(125 \pm 5)$  °С. Данный параметр устанавливался как наименьшее значение из трех величин:

А) максимальной температуры стеклования материала платы (для стеклотекстолита типа FR4 данная температура составляет 125°С, для стеклотекстолита типа FR5, а также ламинатов типа Rogers, Hitachi температура стеклования обычно не превышает 150 °С);

Б) максимально возможной температуры эксплуатации готового изделия;

В) максимально возможной температуры нагрева корпуса самого чувствительного компонента (указывается в ТУ или datasheets).

До 2021 года соединители на данные платы устанавливали ручным монтажом. Плату с соединителями, зафиксированными крайние выводы, на специальных приспособлениях – подставках устанавливали на подогревной столик на расстоянии 11 мм от его поверхности, разогретого до температуры  $(220 \pm 10)$  °С, и выдерживали на столике от 10 до 15 минут. За данное время поверхность платы нагревалась до 90 °С. После этого с помощью паяльника, температура жала которого составляла  $(280 \pm 10)$  °С, выполняли пайку штыревых выводов. Среднее время пайки выводов составляло от 8 до 10 с (в зависимости от того, сколько металлизированных полигонов соединялось с тем или иным отверстием). Температуру жала паяльника, поверхности подогревного столика, а также температуру платы в отверстии контролировали с помощью мультиметра типа APPA305 стандартной термопарой (в случае контроля температуры поверхности платы термопару закрепляли на плате с помощью термостойкого скотча). Общее количество точек пайки, равных количеству запаиваемых штыревых выводов, составляло соответственно 60 шт для адаптерного и 74 шт для коммутаторного блока. Среднее время выполнения операции ручной пайки для одного модуля составляло: 1,2 часа для адаптерного блока и 1,5 часа для коммутаторного блока.

При отработке режима селективной пайки важно было получить качественные паяные соединения, удовлетворяющие установленным выше требованиям, не превышая при этом отработанных годами температурно-временных пределов технологии ручной пайки (максимальной температуры нагрева платы, температур и времени пайки и т.п.). Для наблюдения за изменением температурно-временных характеристик на плате использовали термопрофайлерс 4 термопарами N-K типа. Термопары с помощью термостойкого скотча закрепляли:

- на корпусе паяемого компонента;

- на плате рядом с посадочным местом компонента (примерно на расстоянии 10 мм от зоны воздействия волны);

- в отверстии платы.

Подготовленную многослойную ПП (высушенную в термошкафе при температуре  $(120 \pm 5)$  °С в течение 3 часов) вместе с термопрофайлером устанавливали в направляющие паллеты и активировали написанную ранее программу, в которой заранее были указаны места воздействия на плату. С помощью гербер-файлов определяли нулевую точку на плате, указывали позиции компонентов под установку, выбирали оптимальный путь движения волны, скорость, время флюсования и пайки. После запуска программы выполнялся рабочий цикл: флюсование, нагрев и пайка выводов ИМО, а термопрофайлер в режиме реального времени передавал информацию об изменении температурно-временных характеристик на специализированное удаленное рабочее место, где с помощью специализированного программного обеспечения строились соответствующие графики (см. рисунок 2). После каждого цикла пайки полученные паяные соединения подвергались отмывке и визуально-оптическому и рентгенодефектоскопическому контролю.



**Рис. 2. Фотография экрана монитора с изображением термопрофиля пайки платы адаптерного блока во время циклов обработки, снятого в режиме реального времени**

В результате был подобран режим пайки, полностью удовлетворяющий установленным критериям: температура пайки составляла 285°C (не выше 290 °C), время пайки 1 точки – 2 с (не более 10 с), температура нагрева поверхности платы во время пайки – 117 с (не более 120 с). Качество полученных паяных соединений полностью соответствовало установленным критериям как по данным рентгенодефектоскопического осмотра (см. рисунок 3), так и визуально при 8<sup>x</sup> увеличении: паяные соединения гладкие, блестящие, без пор и посторонних включений, форма соединения скелетная с вогнутыми галтелями.



**Рис. 3. Иллюстративный снимок паяных соединений соединителей на плате коммутаторного блока. Регистрируется полное протекание припоя в отверстия платы в объеме 100%.**

Данная технология была применена при сборке 98 плат адаптерного и 44 плат коммутаторного блоков. Время выполнения операции сократилось с 1,5 часов до 7 минут, дополнительных ремонтных операций не потребовалось из-за высокого качества 100% паяных соединений на всех платах. Сравнительный анализ возможностей двух технологий представлен в таблице 1.

**Таблица 1**

**Сравнительный анализ технологий селективной пайки и ручного монтажа с помощью паяльника и подогревного столика**

<b>Пайка с помощью паяльника</b>	<b>Критерий сравнения</b>	<b>Селективная пайка</b>
Не регламентировано	Габаритные размеры ПП, мм	От 15×15 до 508×508
Не регламентировано	Толщина ПП, мм	От 0.5 до 4
От 0.8 – 2,5 (В допуске от 0.5) *	Высота выступающей части вывода над поверхностью платы со стороны пайки, мм	От 0.8 – 2.0 (В допуске от 0.5)*
Не регламентировано	Максимальная высота корпусов компонентов со стороны пайки	До 60 мм
Нет	Наличие реперных знаков	Да (круги диаметром 1 мм)
Нет	Наличие гербер-файлов слоев ПП	Да
Нет	Наличие технологических полей	Да (10 – 15 мм)
Нет	Ограничение по расположению компонентов и печатного рисунка на краях ПП	3 мм
Нет (Зависит от исполнителя)	Повторяемость результатов	Да (регламентируется программой)
Нет (Зависит от исполнителя)	Стабильность поддержания параметров режимов	Да (регламентируется программой)
Воздух	Рабочая среда	Азот
Температура жала паяльника: (240 – 290) °С, в зависимости от типа ИМО и структуры платы	Температура пайки (припой ПОС61)	Температура мини-волны (260 – 290) °С, в зависимости от типа ИМО и структуры платы
Да Осуществляется дополнительным оборудованием – подогревным столиком, с дополнительной оснасткой. Для пайки припоем ПОС61: -от 60°С до 110°С (тип плат FR4, отечественные и зарубежные ИМО) -от 60°С до 140°С (тип плат FR5, с учетом конструкции ранее смонтированных ЭРИ)	Дополнительный нагрев ПП	Да Осуществляется в едином цикле в одной установке Для пайки припоем ПОС61: -от 60°С до 110°С(тип плат FR4, отечественные и зарубежные ИМО) -от 60°С до 140°С(тип плат FR5, с учетом конструкции ранее смонтированных ЭРИ)
Единичные образцы, серийный выпуск	Тип выпуска продукции	Серийный выпуск (требуется отработка режима пайки и написание программы пайки для конкретной платы)
1,2 – 1,5 часа	Время проведения операции монтажа ИМО	6,5 – 7 минут



\*согласуется с конструктором изделия

Как видно из сравнительных данных, приведенных в таблице 1, технология селективной пайки – частный случай традиционной технологии пайки паяльником. Однако учитывая высокую стоимость оборудования, строго определенные условия его эксплуатации (работа в среде азота определенного давления и чистоты), а также необходимость введения специальных изменений в конструкцию печатных плат (введение технологических полей и реперных знаков, учет высот компонентов на поверхности платы), необходимость предварительной технологической подготовки производства, в виде отработки программ пайки, ограничивает использование в опытном производстве. Свою эффективность селективная технология может реализовать в качестве серийного производства.

#### **Библиографический список**

1. Портал Элинформ. <http://www.elinform.ru> «Основы технологии монтажа в отверстия. Часть I» – 24 июня 2007.
2. Военный стандарт ОСТ В95 2302-82. Аппаратура микроселектронная. Пайка выводов электрорадиоизделий на печатные платы. Типовые технологические процессы. Общие технические требования к качеству паяных соединений; Введ. 01.07.82, 2021. – 88 с.
3. IPC-A-610D Acceptability of Electronic Assemblies – February 2005.
4. Директива RoHS 2011/65/EU
5. ОСТ В 11.073.063. Микросхемы интегральные. Выбор и определение допустимых значений параметров воздействующих технологических факторов при производстве радиоэлектронной аппаратуры на интегральных микросхемах. – Взамен ОСТ II 073.063-76; Введ. 05.09.84.

**ОПТИМИЗАЦИЯ МАРШРУТА СБОРКИ ЭЛЕКТРОННЫХ МОДУЛЕЙ НА ОСНОВЕ  
МНОГОСЛОЙНЫХ КЕРАМИЧЕСКИХ LTCC-ПЛАТ ЗА СЧЕТ ВНЕДРЕНИЯ ПРОЦЕССА  
АВТОМАТИЧЕСКОГО ДОЗИРОВАНИЯ**

**И.П. Чиненкова**

ORCID: 0000-0002-4482-1869 e-mail: irinka13.chip162@gmail.com

**И.О. Усова**

ORCID: 0000-0002-8788-4093 e-mail: usova\_i@mail.ru

**С.В. Кузнецова**

ORCID: 0000-0002-7326-5128 e-mail: svlika@mail.ru

Филиал РФЯЦ-ВНИИЭФ «НИИИС им. Ю.Е. Седакова

В докладе рассматриваются особенности процесса автоматического дозирования припойной пасты в полости многослойных керамических плат. Приводятся сравнительный анализ двух существующих технологий группового нанесения припойной пасты (нанесение через трафарет и автоматическое дозирование), рекомендации по выбору формы отпечатков, количества точек дозирования, а также – оптимизированный технологический маршрут сборки.

**Ключевые слова:** многослойная керамическая плата, технология LTCC, припойная паста, трафаретная печать, дозирование, форма отпечатка припойной пасты, поверхностный монтаж, конвекционная пайка.

**OPTIMIZATION OF THE ASSEMBLY ROUTE OF ELECTRONIC MODULES BASED ON  
MULTILAYER CERAMIC LTCC BOARDS THROUGH THE INTROPRODUCTION OF AN  
AUTOMATIC DOSING PROCESS**

**I.P. Chinenkova, I.O. Usova, S.B. Kuznetsova**

Branch of federal state unitary enterprise «Russian federal nuclear center - All-russian scientific-research institute of experimental physics» «Scientific-research institute of measuring systems  
n.a. Yu.E. Sedakov»

**Purpose:** Optimization of technological processes for manufacturing electronic modules based on multilayer ceramic boards and expanding the possibilities of creating products of fundamentally new designs due to the development of components not only on flat surfaces, but also in cavities.

**Design/methodology/approach:** A new method for applying solder paste to ceramic boards is proposed, which allows not only to improve the quality of the resulting solder paste prints, stability and repeatability, but also to produce fundamentally new designs.

**Findings:** An optimized technological route for assembling an electronic module based on multilayer ceramics has been developed, which makes it possible to install surface-mounted products not only on the surface, but also in cavities of various depths.

**Research limitations/implications:** This article is a continuation of the complex of works on the manufacture of electronic modules based on ceramic LTCC boards.

**Originality/value:** Using the automatic dosing unit makes it possible to apply solder on multi-level ceramic plates in automatic mode, which, in turn, will allow serial assembly of such plates.

**Keywords:** multilayered ceramic plate, Low Temperature Co-fired Ceramic, solder paste, screen printing, dosage, solder paste print shape, surface mounting, convection soldering.

## **Введение**

Операция нанесения припойной пасты – это одна из основных операций при сборочно-монтажном производстве. Причины появления 80% [1] всех дефектов паяных соединений связывают с некачественным выполнением именно этого процесса.

В поверхностном монтаже операция нанесения припойной пасты на контактные площадки (КП) производится двумя основными методами: трафаретной печатью и дозированием. Реализуется ручным способом и с помощью полуавтоматического и автоматического оборудования.

Трафаретная печать является наиболее распространенным методом в серийном производстве, но требует использования дополнительной оснастки – трафарета. Использование автоматических установок трафаретной печати обеспечивает самую высокую производительность и качество печати печатного рисунка любой сложности, выполненного на плоской поверхности (BGA, компоненты с шагом выводов менее 0,5 мм, наличие большого разнообразия компонентов на плате).

В экспериментальном производстве, при несложном печатном рисунке (под компоненты с шагом выводов не более 0,5 мм, не содержащий BGA) чаще всего используют метод дозирования. Большое преимущество дозирования заключается в отсутствии дополнительной оснастки – трафарета. Эту технологию эффективно использовать при ремонтных работах (при нанесении пасты на посадочное место заменяемого компонента), в мелкосерийном производстве, при сборке единичных плат малого размера, при нанесении пасты на неплоские многоуровневые объекты. Именно такими объектами являются многослойные керамические платы (МКП), выполненные по технологии LTCC (Low Temperature Co-fired Ceramic) – технология низкотемпературной совместно обжигаемой керамики.

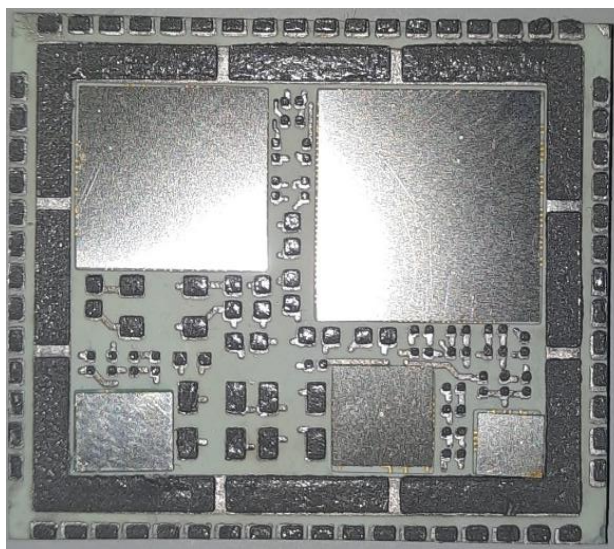
В настоящее время создание МКП и электронных модулей (ЭМ) на их основе [2, 3] являются одним из востребованных продуктов и имеют ряд преимуществ:

- расширенный диапазон механических, тепловых и рабочих нагрузок;
- высокие технические характеристики и надежность;
- вакуумплотность;
- высокая повторяемость параметров;
- компактные размеры и прочность конструкции.

Соответственно, задача оптимизации технологических процессов изготовления ЭМ на основе МКП, уменьшение их себестоимости и расширение возможностей создания изделий принципиально новых конструкций за счет освоения автоматизированного монтажа компонентов является одним из приоритетных направлений работ по производству таких изделий.

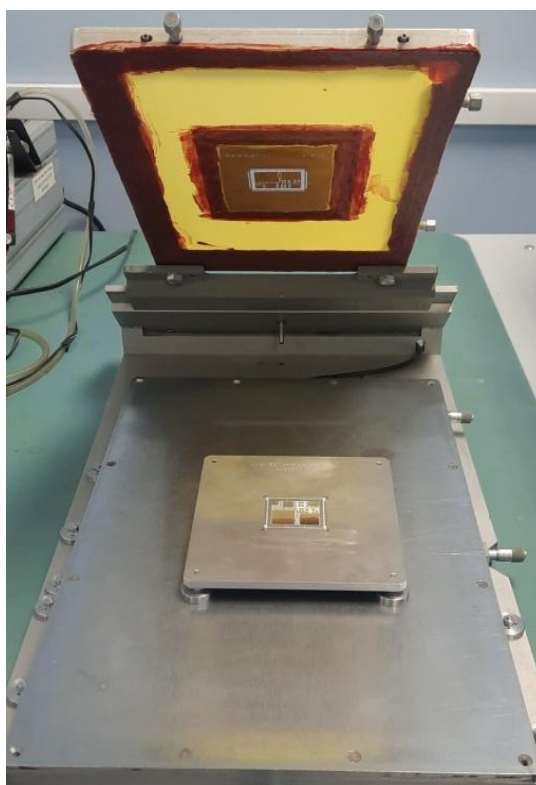
## **Экспериментальная часть**

В 2019 году был проведен ряд экспериментальных работ, результаты которых были отражены в технологическом маршруте [1] и реализованы в нескольких партиях СВЧ-модулей на основе МКП. Из-за малых размеров МКП, невозможности мультиплицирования, а также низкой точности выполнения габаритных размеров и печатного рисунка, сборка данных изделий целиком на автоматизированной линии пока не представляется возможным. Для монтажа поверхностно-монтируемых изделий (ПМИ) на припойную пасту был использован метод трафаретной печати (рис. 1), реализуемый на ручном устройстве.



**Рис. 1. Фотография LTСС-платы с отпечатками припойной пасты, нанесенными через трафарет**

Нанесение пасты на контактные поверхности МКП платы электронного модуля, содержащие посадочные места под ПМИ и под титановую обечайку прямоугольной формы, осуществлялось через комбинированный сетчато-латунный трафарет толщиной 0,15 мм на установке трафаретной печати начального уровня модели 903002 ф. Fritch (рис. 2). Собираемые стороны модуля были спроектированы таким образом, чтобы ПМИ располагались на плоских поверхностях платы [1]. Объемы пасты, нанесенные на КП, были определены размерами апертур трафарета и требованиями [4], и позволили получить качественные паяные соединения, соответствующие требованиям [5]. Партии модулей успешно выдержали все необходимые испытания и подтвердили свою работоспособность.



**Рис. 2. Фотография LTСС-платы в специальном приспособлении-кондукторе на рабочем поле установки 903002**

Необходимо отметить, что операция нанесения пасты, выполненная на установке ручной трафаретной печати, была достаточно затратной по времени и средствам, вложенным в изготовление необходимой оснастки (проектирование и изготовление сетчато-латунного трафарета, его натяжение на раму, изготовление кондуктора), требовала повышенного внимания и аккуратности со стороны исполнителя (совмещение апертур трафарета и КП платы выполнялось вручную), обладала низкой повторяемостью. Нестабильная усадка керамических материалов от партии к партии приводила к тому, что оснастку необходимо было изготавливать под каждую партию МКП индивидуально.

В 2021 году был разработан новый прибор, в основу которого легла 28-слойная керамическая плата, в которой ПМИ располагались уже не только на внешних поверхностях, но и во внутренних колодцах на различной глубине. Миниатюрные размеры полостей и плотность печатного рисунка не обеспечивали доступ жала паяльника к паяемым внутренним поверхностям. Размещение локального трафарета в полостях оказалось также нереализуемым. Данные обстоятельства привели к необходимости разработки и освоения нового технологического процесса – автоматического дозирования припойной пасты.

При разработке данной технологии в приложении к особенностям печатного рисунка МКП были определены:

- оптимальный диаметр и конструкция иглы (0,51 мм для КП посадочных мест под чип 0805 и более, а также под КП прямоугольной замкнутой полосообразной формы – обечайку, 0,33 мм для КП посадочных мест под чип 0603, 0402, указанные на рис. 3);
- типы возможных отпечатков (однократные и многократные точечные отпечатки, однократные и двукратные линейные отпечатки) в расчете на габаритные размеры КП посадочных мест ПМИ;
- параметры режима нанесения точечных и линейных отпечатков (время нанесения, скорость перемещения иглы и зазор между печатной платой (ПП) и наконечником иглы).

Диаметр и конструкцию иглы определяли из площади КП. При отработке и определении доз припоя исходили из того, что фактический объем припойной пасты, нанесенной на конкретную КП должен быть сопоставим с аналогичным объемом пасты, формируемым на данном КП в результате трафаретной печати. Дозы пасты должны формировать паяные соединения, полностью удовлетворяющие требованиям международного стандарта [5]. Экспериментально было определено, что для КП под компоненты типоразмера 0603, 0805 достаточно однократных точечных отпечатков, для остальных КП необходимо использовать либо двукратные точечные отпечатки, либо однократные линейные отпечатки.

Параметры режима нанесения отпечатков были получены в процессе отработки, изменяя время нанесения (для точечных отпечатков) от 0,08 с до 0,2 с и скорость перемещения (для линейных отпечатков) – от 2 мм/с до 5 мм/с. Зазор между ПП и наконечником иглы для всех отпечатков установили примерно равный 0,127 мм.

Экспериментальную отработку и последующие сборки образцов производили на установке автоматического дозирования DispenseMate 595. Для нанесения использовали припойные пасты на основе сплава Sn62Pb36Ag2 (припойная паста производства МЭФ «ОНИКС» г. Ярославль, марки 7141 типа 3 и ее импортный аналог SNQ90 типа 5, производства Indium Corporation of America, США).

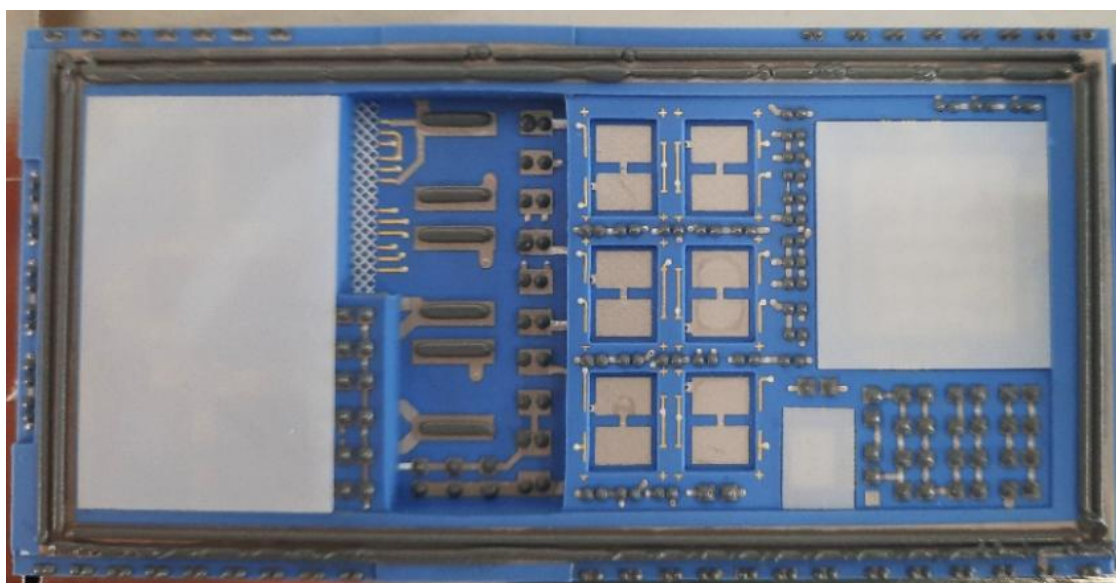
Последовательность нанесения отпечатков была заложена в управляющей программе (УП), в которой указывали координаты центров КП, типы и параметры возможных отпечатков, безопасную высоту опускания иглы (для нанесения пасты не только на внешние КП, но и во внутренние колодцы платы). Фотография фрагмента управляющей программы показана на рис. 3.

Необходимо отметить факт, что при изготовлении МКП происходит их усадка в соответствии с коэффициентами, указанными в характеристиках материала плат. При обжиге МКП происходит разная усадка от партии к партии. Интервал колебания габаритных размеров МКП составляет  $\pm 300$  мкм (согласно допуску), внутренних полостей МКП – до 100 мкм, что существенно усложняет работу по нанесению пасты автоматическим дозированием, увеличивая время нанесения пасты из-за индивидуальной коррекции УП под каждую плату.

Program: keramika_LTCC		Fluid1: NIDLE 25 T5 NO
1	SET ACTIVE VALVE: 1	
2	DOT: 1, {24.002, 18.553}	
3	DOT: 1, {25.586, 18.571}	
4	DOT: 1, {23.967, 16.244}	
5	DOT: 1, {25.535, 16.204}	
6	DOT: 1, {23.970, 13.961}	
7	DOT: 1, {25.589, 14.052}	
8	DOT: 1, {24.075, 11.478}	
9	DOT: 1, {25.583, 11.568}	
10	DOT: 1, {24.094, 8.889}	
11	DOT: 1, {25.642, 8.835}	
12	DOT: 1, {24.112, 6.391}	
13	DOT: 1, {25.714, 6.446}	
14	FIND SUBSTRATE HEIGHT: {24.001, 16.254},0,0	
15	FIND SUBSTRATE HEIGHT: {29.175, 32.318},0,0	
16	LINE: 1, Start:{29.445, 32.517}, End:{33.776, 32.246}	
17	LINE: 1, Start:{28.936, 26.850}, End:{33.756, 26.923}	
18	LINE: 1, Start:{28.920, 24.146}, End:{33.684, 24.020}	
19	LINE: 1, Start:{29.220, 18.785}, End:{33.786, 18.788}	
20	LINE: 1, Start:{28.891, 15.783}, End:{33.723, 15.966}	
21	LINE: 1, Start:{29.287, 10.612}, End:{33.909, 10.432}	
22	LINE: 1, Start:{36.635, 32.286}, End:{38.274, 32.413}	
23	LINE: 1, Start:{36.509, 29.426}, End:{38.201, 29.481}	
24	LINE: 1, Start:{36.509, 26.566}, End:{38.093, 26.603}	
25	LINE: 1, Start:{36.383, 23.670}, End:{37.913, 23.815}	
26	LINE: 1, Start:{36.329, 20.955}, End:{37.985, 20.955}	
27	LINE: 1, Start:{36.293, 17.950}, End:{37.949, 17.968}	
28	LINE: 1, Start:{36.473, 14.927}, End:{37.985, 15.072}	
29	LINE: 1, Start:{36.527, 12.211}, End:{38.237, 12.248}	
30	LINE: 1, Start:{36.437, 9.351}, End:{38.111, 9.424}	
31	DOT: 1, {29.023, 7.795}	
32	DOT: 1, {29.274, 6.310}	
33	DOT: 1, {31.488, 7.849}	
34	DOT: 1, {31.704, 6.165}	
35	DOT: 1, {33.971, 7.939}	
36	DOT: 1, {33.989, 6.111}	
37	LINE: 1, Start:{36.490, 6.473}, End:{37.948, 6.582}	

**Рис. 3. Фотография фрагмента управляющей программы «keramika\_LTCC» для нанесения отпечатков припойной пасты на LTCC-плату**

Визуально-оптическая оценка качества отпечатков припойной пасты показала, что полученные отпечатки имели круглую форму, занимали более 90% их площади, смещения отпечатков относительно установленных программой положений – отсутствовали. Фотография МКП с нанесенной пастой на КП, представлена на рис. 4.



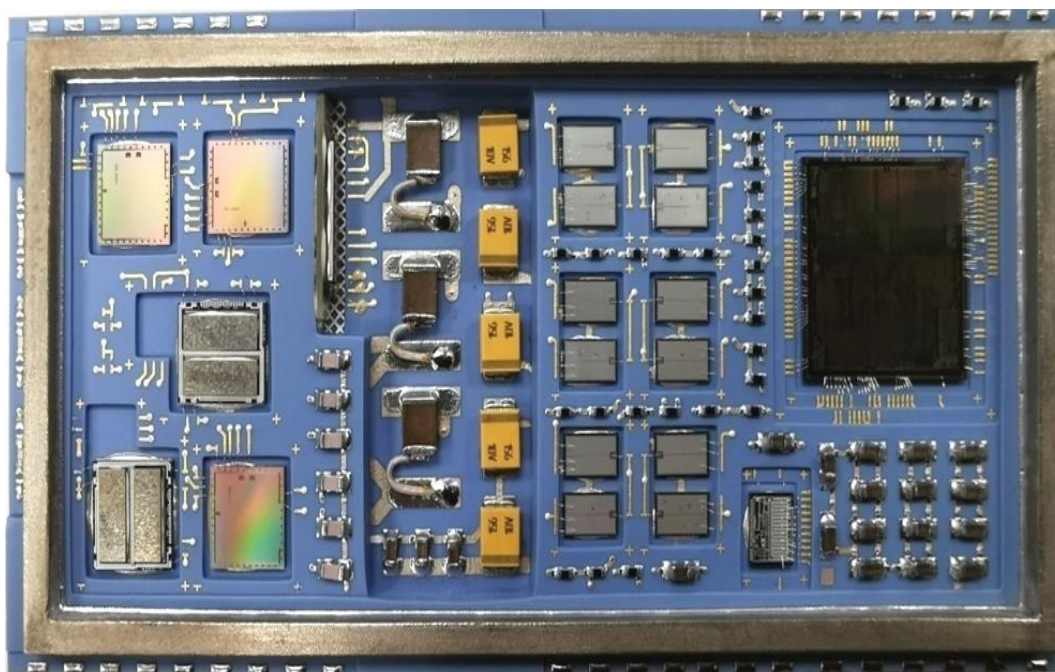
**Рис. 4. Фотография LTCC-платы с отпечатками припойной пасты, нанесенными с помощью установки дозирования**



Далее на МКП по вновь разработанному маршруту были установлены и запаяны в печи конвекционного оплавления припоя требуемые ПМИ. В данном случае были установлены в полости керамические вкладыши, чтобы обеспечить защиту посадочных мест кристаллов от воздействия материалов поверхностного монтажа. Это необходимо для того, чтобы золотые контактные поверхности в полостях, предназначенных для последующей разварки установленных кристаллов, были пригодны к сварочным операциям после прохождения операций пайки – на них не должно быть следов припоя, флюса, механических повреждений в виде царапин. Для защиты золотых КП от случайного попадания шариков припоя из-за кипения флюсующей составляющей припойной пасты во время прохождения модуля по конвейеру конвекционной печи используются керамические (поликоровые) или стальные вкладыши толщиной 200 мкм, повторяющие форму монтажных колодцев.

После пайки МКП с припаянными ПМИ и установленными вкладышами, помещали в специализированную оснастку (пластиковый ячеистый короб) и подвергали механизированной отмывке. Вкладыши извлекали из полостей только после отмывки.

Качество паяных соединений полностью удовлетворяло требованиям [5]. Фотография собранного электронного модуля на основе МКП представлена на рис. 5. Работоспособность модулей была подтверждена положительными результатами испытаний.



**Рис. 5. Фотография собранного электронного модуля на основе ЛТСС-платы**

Сравнительный анализ двух методов нанесения пасты, используемых в процессе изготовления модулей на основе МКП, приведены в табл. 1.

Данные содержащиеся в табл.1, позволяют сделать однозначный вывод о том, что при равном качестве получаемых паяных соединений наиболее оптимальным решением для керамических плат признан метод автоматического дозирования, вне зависимости от особенностей конструктива МКП (с плоскими монтируемыми сторонами или многоуровневые конструкции). Метод трафаретной печати пригоден только для плат, где ПМИ располагаются на плоских монтируемых сторонах. Выбор типа припойной пасты, иглы и типов отпечатков определяется сложностью печатного рисунка МКП.

Таблица 1

## Сравнение методов нанесения припойной пасты на LTCC-платы

Дозирование	Характеристика	Трафаретная печать
Не требуются	Дополнительные приспособления	Трафарет и кондуктор на каждую партию
Да (для нанесения пасты)	Необходимость создания УП	Да (для изготовления трафарета)
Оператор	Специалист, создающий УП	Конструктор
Автоматический	Тип нанесения	Ручной
Нет	Наличие дополнительной операции по доработке отпечатков	Да (для прямоугольных апертур под конструкционные элементы типа обечайки)
- зазор между платой и наконечником иглы; - диаметр и конструкция иглы; - время и скорость нанесения - тип отпечатка, определяемый в соответствии с размерами КП	Параметры, определяющие дозу припойной пасты на КП	- размеры апертуры трафарета, определяемый в соответствии с размерами КП; - толщина полотна трафарета; - способ изготовления трафарета; - скорость движения ракеля по полотну трафарета; - угол наклона и давление ракеля; - величина зазора между платой и трафаретом; - высота отделения трафарета от платы; - скорость отделения трафарета от платы; - чистота апертур
На плоские поверхности и на внутренние полости	Области нанесения пасты	Только на плоские поверхности
Да	Возможность двустороннего монтажа	Да
Да	Воспроизводимость дозы пасты	Нет (зависит от человеческого фактора)
В пределах одной партии МКП из-за усадки керамических материалов	Повторяемость процесса	В пределах одной партии МКП из-за усадки керамических материалов
Есть (зависит от программы)	Стабильность	Нет (зависит от человеческого фактора)
Да	Качество полученных паяных соединений соответствует требованиям [5]	Да
Да (только для печатного рисунка, не содержащего посадочные места под чип 0603 и менее, под микросхемы в корпусах QFN, QFP с шагом выводов 0,8 мм и менее, BGA с диаметром шариковых выводов 0,8 мм и менее)	Использование отечественной припойной пасты типа 3	Да (только для печатного рисунка, не содержащего посадочные места под чип 0201 и менее, под микросхемы в корпусах QFN, QFP с шагом выводов 0,5 мм и менее, BGA с диаметром шариковых выводов 0,45 мм и менее)
20 минут (с учетом предварительной калибровки и самого нанесения)	Время протекания процесса нанесения	2 часа (с учетом ручного совмещения, дополнительной операции ручного добавления пасты, исправления дефектов, перенанесений)



## **Выводы**

В результате работы было выявлено, что метод автоматического дозирования наиболее оптимальное решение для нанесения припойных паст на керамические платы. Был разработан оптимизированный технологический маршрут сборки электронного модуля на основе МКП, который позволяет устанавливать ПМИ не только на плоскую поверхность, но и полости различной глубины. Приведены рекомендации по выбору форм отпечатка, количества точек дозирования. Написана управляющая программа для данного электронного модуля. Приведены параметры режимов нанесения, которые были получены в процессе отработки управляющей программы. При выполнении всех этих требований были собраны 5 работоспособных электронных модулей на основе МКП.

## **Библиографический список**

1. **Усова, И.О., Кондрашов, Д.А., Краснова, Н.Н.** Исследования и разработка технологических маршрутов сборки электронных модулей на основе многослойных керамических плат из отечественной системы материалов // Информационные системы и технологии. – 2019. – С.217 – 222.
2. **Симин, А., Холодняк, Д., Вендик, И.** Многослойные интегральные схемы сверхвысоких частот на основе керамики с низкой температурой обжига // Компоненты и технологии. – 2005. – №5. – С.190 – 196.
3. **Ляпин, Л.В., Сытилин, С.Н., Осипов, А.В.** Отечественная система низкотемпературной керамики LTCC СКМ // Материалы межотраслевого семинара «Вопросы освоения LTCC-технологии на Российских предприятиях, разработки и изготовления на ее основе СВЧ аналоговой и цифровой РЭА». – Н. Новгород: ОАО «ФНПЦ «ННИИРТ». – 2009.
4. IPC-7525 Stencil Desing Guidelines – May 2000.
5. IPC-A-610D Acceptability of Electronic Assemblies – February 2005.

## ИЗМЕРЕНИЕ ВОЛНОВОГО СОПРОТИВЛЕНИЯ МИКРОПОЛОСКОВОЙ ЛИНИИ

Е.А. Лупанова

ORCID: 0000-0001-6389-9939 e-mail: ealupanova@yandex.ru

С.М. Никулин

ORCID: 0000-0003-3906-8371 e-mail: nikulin-serg2006@yandex.ru

Нижегородский государственный технический университет им. Р.Е. Алексеева

В докладе предложен способ измерения волнового сопротивления микрополосковой линии передачи векторным анализатором цепей. В диапазоне от 0,02 до 2,0 ГГц волновое сопротивление определяют из результатов измерений S-параметров T и L-соединений коаксиально-полосковых переходов с отрезками короткой и электрически длинной линии. Частотную зависимость волнового сопротивления в диапазоне от 0,02 до 26,50 ГГц определяют с помощью широкополосного TRL-метода и схемной модели полосковой линии в системе автоматизированного проектирования AWR. В основу предлагаемого решения положена концепция стандарта Битти. Актуальность решения задачи показана на примере идентификации собственных параметров чип резисторов.

**Ключевые слова:** микрополосковая линия, волновое сопротивление, эффективная диэлектрическая проницаемость, S-параметры, векторный анализатор цепей, LTR -метод, стандарт Битти, система автоматизированного проектирования, схемная модель, чип резистор

## MEASUREMENT OF THE WAVE RESISTANCE OF A MICROSTRIP LINE

E.A. Lupanova, S.M. Nikulin

Nizhniy Novgorod state technical university n.a.R.E. Alekseev

**Purpose:** The article proposes a method for measuring the wave impedance of a microstrip transmission line with a vector network analyzer.

**Design/methodology/approach:** In the range from 0.02 to 2.0 GHz, the wave impedance is determined from the results of measurements of the S-parameters T and L-connections of coaxial-strip junctions with segments of a short and electrically long line. The frequency dependence of the wave impedance in the range from 0.02 to 26.50 GHz is determined using the broadband TRL method and the circuit model of strip lines in the AWR computer-aided design system. The proposed solution is based on the concept of the Beatty standard.

**Findings:** New method for determining the wave impedance of a microstrip line from the measurement results is proposed.

**Research limitations/implications:** Research to determine the wave impedance of microstrip lines continues.

**Originality/value:** The relevance of solving the problem is shown by the example of identifying the intrinsic parameters of chip resistors.

**Keywords:** microstrip line, wave impedance, effective permittivity, S-parameters, vector network analyzer, LTR method, Beatty standard, computer-aided design system, circuit model, chip resistor

## Введение

Волновое сопротивление микрополосковой линии определяется ее конструкцией, геометрией поперечного сечения, относительной диэлектрической проницаемостью изоляционного основания и другими факторами, которые не поддаются точному оцениванию и аналитическому моделированию. В работе [1] предложена аналитическая модель волнового сопротивления, параметры которой в принципе можно установить с помощью СВЧ измерений, однако вопрос о достоверности получаемых оценок остается открытым ввиду того, что поиск ряда параметров предлагается проводить на основе измерительной информации, которая не поддается верификации и стандартизации.

Цель исследования состоит в поиске альтернативных решений, которые имеют логическую основу, доказательную базу и содержат информацию о рассеянии получаемых оценок.

### Измерение волнового сопротивления в диапазоне 0.02 – 2.0 ГГц

Волновое сопротивление микрополосковой линии  $\dot{Z}_B$  является комплексной величиной в силу конечной проводимости проводников и отличного от нуля тангенса угла диэлектрических потерь [2]:

$$\dot{Z}_B = Z_B - i \frac{1}{2} Z_B (R/\omega L - G/\omega C). \quad (1)$$

Здесь  $Z_B = \sqrt{L/C}$  - действительная величина  $\dot{Z}_B$ ,  $L, C, R, G$  - соответственно индуктивность, емкость, сопротивление и проводимость лестничной эквивалентной схемы линии с распределенными параметрами. Ввиду малой величины второго слагаемого емкостной характер  $\dot{Z}_B$  проявляется лишь в низкочастотной части СВЧ диапазона. В области контакта коаксиально-полоскового перехода с полосковой линией возникает возмущение электромагнитного поля, которое можно интерпретировать как неоднородность индуктивного характера. Причем, как показали экспериментальные исследования, существует одна или несколько дискретных частот, при которых наблюдается полная компенсация емкостной составляющей волнового сопротивления индуктивной неоднородностью области контакта. Этот эффект позволяет выполнить измерение  $\dot{Z}_B$  в низкочастотной части СВЧ диапазона векторным анализатором цепей.

С этой целью следует измерить анализатором цепей S-параметры соединения двух коаксиально-полосковых переходов с отрезком микрополосковой линии и перенести результаты к областям контакта центральных проводников коаксиальных линий с полосковым проводником [1]. Затем найти отношения волновых сопротивлений  $r_{a,b} = \dot{Z}_B^{a,b}/50$  и коэффициенты передачи  $t_{a,b}$  линии с помощью быстросходящихся итерационных процедур вычислений:

$$\begin{aligned} r_a(f) &= (1 + s_{11}(f) - s_{21}(f)t_a(f)) / (1 - s_{11}(f) - s_{21}(f)t_a(f)), \\ t_a(f) &= s_{21}(f)(r_a(f) + 1) \left[ 1 - ((r_a(f) - 1)t_a(f) / (r_a(f) + 1))^2 \right] / (4r_a(f)), \\ r_b(f) &= (1 + s_{22}(f) - s_{12}(f)t_b(f)) / (1 - s_{22}(f) - s_{12}(f)t_b(f)), \\ t_b(f) &= s_{12}(f)(r_b(f) + 1) \left[ 1 - ((r_b(f) - 1)t_b(f) / (r_b(f) + 1))^2 \right] / (4r_b(f)). \end{aligned} \quad (2)$$

На первом шаге итерационных процедур (2)  $t_a$  и  $t_b$  следует принять равными нулю. Волновое сопротивление  $\dot{Z}_B$  определяют на частотах  $f_i$  и  $f_k$ , из соотношений:

$$Z_B^a(f_i) = 50 \operatorname{Re}(r_a(f_i)), \quad Z_B^b(f_k) = 50 \operatorname{Re}(r_b(f_k)), \quad (3)$$

при которых мнимые части  $\dot{r}(f)$  приобретают значения, равные нулю  $-\operatorname{Im}(\dot{r}_a(f_i)) = 0$  и  $\operatorname{Im}(\dot{r}_b(f_k)) = 0$ .

Численные оценки (3) волновых сопротивлений  $Z_B^a(f_i)$  и  $Z_B^b(f_k)$ , показанные на рисунке

1, были найдены из результатов измерений S-параметров соединений коаксиально-полосковых переходов 292-04А-5 с экранированными микрополосковыми линиями. Линии длиной 30 мм и 120 мм были изготовлены из фольгированного материала RO4003С (ROGERS Corp., США). Оценки волнового сопротивления, полученные по S-параметрам L-соединения, имеют разброс значений относительно средней величины:  $46,2 \pm 1$  Ом. Вся совокупность полученных оценок показывает тенденцию к уменьшению  $Z_B$  с увеличением частоты, причем зеленые и синие точки как бы лежат на воображаемой дугообразной линии.

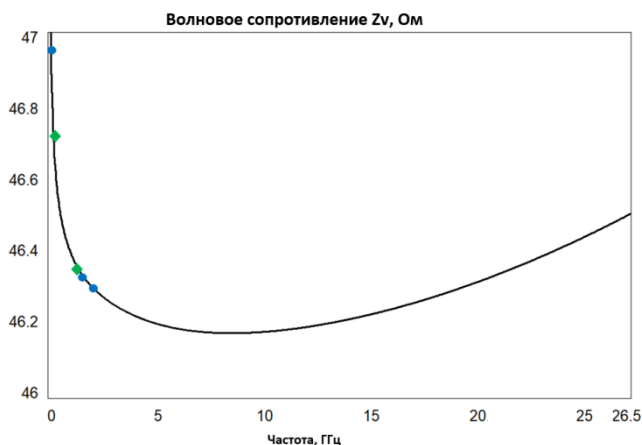


Рис. 1. Волновое сопротивление (зеленые точки - линия 120 мм, синие точки – линия 30 мм).

Альтернативным способом измерения волнового сопротивления микрополосковой линии анализатором цепейв низкочастотной части СВЧ диапазона является LTR-метод [1], в котором оценку волновых сопротивлений  $Z_B^a$  и  $Z_B^b$  предлагается получать из соотношений:

$$\begin{aligned} Z_B^a(f_1) &= 50 \left( 1 - |s_{11}^a(f_1)|^2 \right) / \left( 1 - 2 \operatorname{Re}(s_{11}^a(f_1)) + |s_{11}^a(f_1)|^2 \right), \\ Z_B^b(f_2) &= 50 \left( 1 - |s_{11}^b(f_2)|^2 \right) / \left( 1 - 2 \operatorname{Re}(s_{11}^b(f_2)) + |s_{11}^b(f_2)|^2 \right), \end{aligned} \quad (4)$$

в которых коэффициенты отражения  $s_{11}^a(f)$  и  $s_{11}^b(f)$  со стороны коаксиальных портов искажающих адаптеров А и В вначале получают на массиве базовых частотных точек, а затем в интервалах между ними аппроксимируют сплайнами Акимы. Частоты  $f_1$  и  $f_2$  получают из условия:

$$\begin{aligned} X_B^a(f_1) &= 100 \operatorname{Im}(s_{11}^a(f_1)) / \left( 1 - 2 \operatorname{Re}(s_{11}^a(f_1)) + |s_{11}^a(f_1)|^2 \right) = 0, \\ X_B^b(f_2) &= 100 \operatorname{Im}(s_{11}^b(f_2)) / \left( 1 - 2 \operatorname{Re}(s_{11}^b(f_2)) + |s_{11}^b(f_2)|^2 \right) = 0, \end{aligned} \quad (5)$$

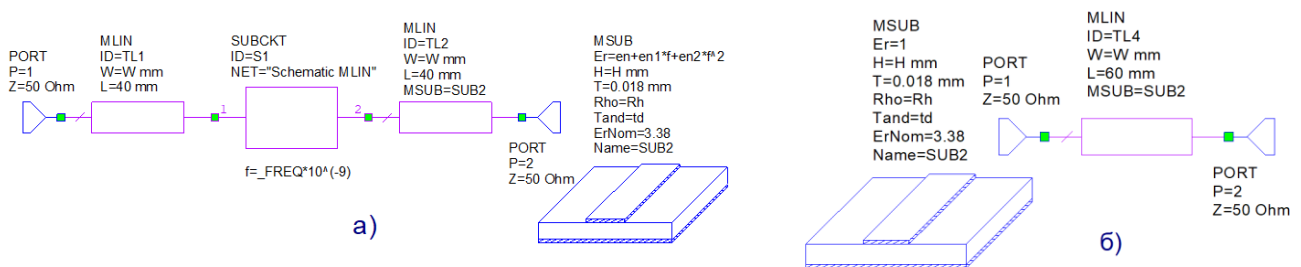
В результате проведенного исследования можно предположить, что оценки волнового сопротивления в низкочастотной области СВЧ диапазона по результатам измерения S-параметров короткого отрезка микрополосковой линии и LTR-методом дают лучшие результаты по сравнению с измерениями L-соединения.

### Определение частотной зависимости волнового сопротивления методом стандарта Битти

Ключевое значение в определении волнового сопротивления микрополосковой линии в СВЧ диапазоне имеют ее относительная эффективная диэлектрическая проницаемость  $\epsilon_{eff}(f)$  и коэффициент затухания  $\alpha(f)$ . В [1] показано, что максимальную точность в решении этой задачи удается получить LTR-методом по сравнению с известным дифференциально-фазовым методом [3]. Частотные зависимости  $\epsilon_{eff}(f)$  и  $\alpha(f)$  позволяют получить средствами оптимизации в САПР AWR схемную модель микрополосковой линии MLIN, адекватную ее физическому прототипу. В результате построения схемного аналога устанавливают аналитическую частотную зависимость относительной диэлектрической проницаемости  $\operatorname{Er}(f)$  изоляционного основания и параметр

$Rh_0 > 1$ , задающий относительно меди сопротивление пленки проводника. Полученная зависимость не столько характеризует истинную величину диэлектрической проницаемости изоляционного основания, сколько моделирует влияние конструктивных особенностей объекта измерений, таких, например, как наличие экранированных слоев на верхней стороне изоляционного основания, зазор между полосковым проводником и экраном, диаметр металлизированных отверстий и расстояние между ними.

Частотную зависимость  $\dot{Z}_B(f)$  предлагается определять методом стандарта Битти [4]. Для микрополосковой линии таким стандартом является отрезок воздушной линии, без изоляционного основания. Техническая реализация микрополоскового стандарта не представляется возможной. По этой причине предлагается воспользоваться средствами автоматизированного проектирования NIAWRDesignEnvironment. На рисунке 2а показана схемная модель микрополоскового стандарта Битти. Между двумя отрезками микрополосковых линий произвольной длины, например, 40 мм, включают блок SUBCKT, моделирующий воздушную полосковую линию длиной L, например L = 60 мм, как показано на рисунке 2б.



**Рис. 2. Трехсекционная схемная модель из отрезков микрополосковых линий передачи**

Выполнив экстракцию портов 1 и 2 вместе с отрезками линий длиной 40 мм Экспресс - методом [5], получают S-параметры воздушной полосковой линии относительно волнового сопротивления  $\dot{Z}_B$ . Отношение волновых сопротивлений воздушной и микрополосковой линий  $r_a(f) = r_b(f) = r(f)$  находят итерационной процедурой (1), а частотную зависимость  $\dot{Z}_B(f)$ , получают из соотношения:

$$\dot{Z}_B(f) = \frac{50}{r(f)} \frac{(1 + S_{11}(f))}{(1 - S_{11}(f))}, \quad (6)$$

в котором коэффициент отражения  $S_{11}(f)$  получают с помощью схемной модели электрически длинной полосковой линии с воздушным диэлектриком.

Волновое сопротивление  $Z \approx 50$  Ом портов 1 и 2 на этой схеме следует выбрать таким, чтобы график частотной зависимости  $\dot{Z}_B(f)$  (линия черного цвета) проходил максимально близко к двум зеленым и двум синим точкам, как показано на рисунке 1. Существенно, что найденную частотную зависимость  $\dot{Z}_B(f)$  можно представить аналитическим выражением, предложенным в [1]:

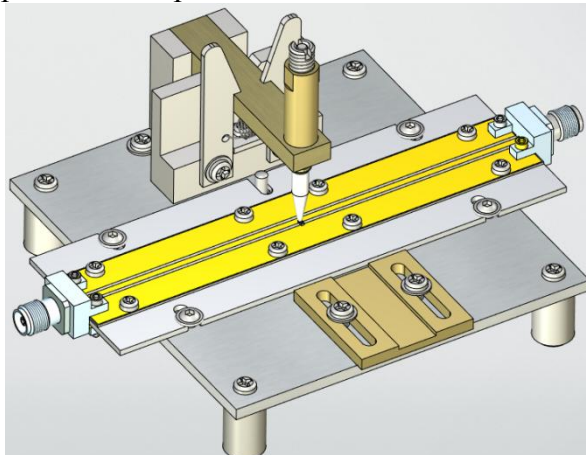
$$\dot{Z}_B(f) = Z_B(f_0) \left\{ q(f) - [q(f) - 1] \frac{\epsilon_r(f_0) - \epsilon_{eff}(f)}{\epsilon_r(f_0) - \epsilon_{eff}(f_0)} \right\}, \quad (7)$$

в котором  $Z_B(f_0) = 46,62$  Ом,  $\epsilon_r(f_0) = 3,39$ ,  $\epsilon_{eff}(f_0) = 2,85$ ,  $q(f) = 1,067 + 0,0011 f / f_k$ ,  $f_k = 26,5$  ГГц.

### Идентификация параметров схемных моделей чип резисторов

Актуальность измерения волнового сопротивления микрополосковой линии подтверждена на примере идентификации параметров схемных моделей чип резисторов трех номиналов 50, 100 и 270 Ом. Измерение S-параметров относительно физических границ объектов выполнялось в контактном устройстве, показанном на рисунке 3. Резистор устанавливался между центральным

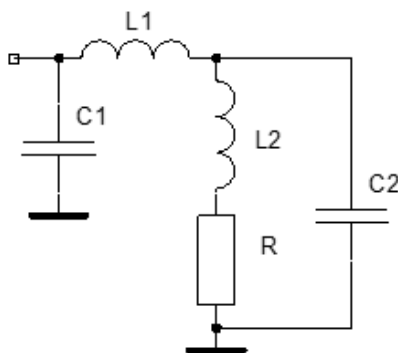
полосковым проводником и экраном при таком подключении контактные площадки отсутствуют, что позволяет определить собственные параметры контролируемого объекта. Экстракция коаксиально-полосковых переходов и отрезков полосковых линий выполнена LTR-методом.



**Рис.3. Оснастка для измерения S-параметров электронных компонентов**

Коэффициенты передачи  $S_{21}$  и  $S_{12}$  инвариантны к положению объектов в оснастке. Влияние погрешности позиционирования на коэффициенты отражения  $S_{11}$  и  $S_{22}$  устранялась заменой этих параметров их полу суммой. Схемная модель резистора показана на рисунке 4. Собственные параметры схемных моделей контролируемых объектов определялись в двух вариантах:

- в предположении о значении волнового сопротивления полосковой линии, равном 50 Ом;
- с фактической величиной волнового сопротивления  $Z_B(f)$ .



**Рис.4. Схемная модель резистора**

Идентификация собственных параметров чип резисторов выполнялась в САПР AWR минимизацией целевой функции Симплекс-методом. Результаты приведены в таблице 1.

**Таблица 1.**

**Параметры схемной модели чип резисторов**

Волновоесопротивление полосковой линии	$50/Z_B(f)$		
	50	100	270
Номинальное сопротивление резистора, Ом	50	100	270
Сопротивление, $R$ , Ом	55,27/51,38	107,75/100,16	286,07/265,92
Емкость $C1$ , пФ	-0,013/-0,021	-0,066/-0,067	-0,014/-0,013
Емкость $C2$ , пФ	0,03/0,04	0,036/0,041	0,022/0,022
Индуктивность $L1$ , нГн	-0,27/-0,20	-0,009/-0,025	-0,216/-0,214
Индуктивность $L2$ , нГн	0,61/0,52	-0,45/-0,30	0,89/0,83

Полученные результаты показывают, что оценки сопротивлений  $R$  для резисторов трех

номиналов с использованием  $Z_B(f)$  оказались более близкими к ожидаемым значениям, чем в случае волнового сопротивления полосковой линии, равного стандартной величине 50 Ом.

### Заключение

Подводя итог выполненного исследования, отметим, что предлагаемый способ измерения волнового сопротивления микрополосковой линии векторным анализатором цепей повышает степень достоверности получаемой информации об электромагнитных параметрах электронных компонентов. По сравнению с зондовыми станциями такое решение представляется более корректным, поскольку воспроизводит условия диагностики объектов, близкие к реальным условиям. Тем не менее, мы понимаем, что обсуждаемая задача все еще далека от окончательного решения ввиду уникального взаимодействия электронных компонентов с реальной линией передачи, в которой они используются. Следует понимать, что даже если удастся исключить окружающие цепи, коаксиально-полосковые переходы, отрезки линий и контактные площадки, схемные модели электронных компонентов в определенной степени несвободны от той линии, в которой проводились испытания. Особенно чувствительны к используемой линии реактивные параметры индуктивности и емкости.

### Библиографический список

1. **Лупанова, Е. А.** Метод определения собственных параметров полосковых линий передачи /Лупанова Е. А., Никулин С. М. // Измерительная техника. – 2021. – №. 5. – С. 47-52.
2. **Wadell, B. C.** Transmission line design handbook. – Artech House Microwave Library, 1991.
3. **Das, N. K.** Two methods for the measurement of substrate dielectric constant / Das N. K., Voda S. M., Pozar D. M. //IEEE Transactions on Microwave Theory and Techniques. – 1987. – Т. 35. – №. 7. – С. 636-642.
4. **Дансмор, Д. П.** Настольная книга инженера. Измерения параметров СВЧ-устройств с использованием передовых методик векторного анализа цепей //М.: ТЕХНОСФЕРА. – 2019.
5. **Евсеев, В. И.** LR и экспресс-метод контроля параметров электронных компонентов в полосковых линиях передачи / Евсеев В. И., Никулин С. М.//Электроника и микроэлектроника СВЧ. – 2020. – Т. 1. – №. 1. – С. 18-22.
6. **Фельдштейн, А. Л.** Справочник по элементам волноводной техники. /Фельдштейн А. Л., Явич Л. Р., Смирнов В. П. // Сов. радио, 1963.

**ИЗМЕРЕНИЕ БИСТАТИЧЕСКОГО КОЭФФИЦИЕНТА ОТРАЖЕНИЯ  
ОСТРОНАПРАВЛЕННЫМИ АНТЕННАМИ НА ФОНЕ МЕШАЮЩИХ ПОМЕХ**

**Е.А. Лупанова**

ORCID: 0000-0001-6389-9939 e-mail: ealupanova@yandex.ru

**С.М. Никулин**

ORCID: 0000-0003-3906-8371 e-mail: nikulin-serg2006@yandex.ru

Нижегородский государственный технический университет им. Р.Е. Алексеева

В настоящий момент измерения бистатического коэффициента отражения требуют наличия безэховой камеры, обработка сигнала проводится методом оконной фильтрации с прямым и обратным преобразованием Фурье, что неизбежно приводит к возникновению эффекта Гиббса. В данной работе предлагается принципиально новый метод измерения бистатического коэффициента отражения на фоне мешающих помех. Измерения проводятся в обычном лабораторном помещении, при обработке сигнала используется принципиально новый метод выявления измеряемой величины на фоне мешающих сигналов, порожденных непрямыми трассами электромагнитных волн в обычном лабораторном помещении. В качестве объекта измерения использовались плоские образцы фторопласта-4.

**Ключевые слова:** бистатический коэффициент отражения, остронаправленные антенны, измерения на фоне мешающих сигналов

**BISTATIC REFLECTION MEASUREMENT POINTED ANTENNA  
IN THE BACKGROUND OF INTERFERENCE**

**E.A. Lupanova, S.M. Nikulin**

Nizhniy Novgorod state technical university n.a.R.E. Alekseev

**Purpose:** The paper proposes a fundamentally new method for measuring the bistatic reflection coefficient against the background of interfering noise.

**Design/methodology/approach:** Currently, measurements of the bistatic reflectance require an anechoic chamber. Signal processing is carried out by window filtering with forward and inverse Fourier transforms, which inevitably leads to the Gibbs effect. In this work, it is proposed to measure in an ordinary laboratory room, while signal processing uses a fundamentally new method for detecting the measured value against the background of interfering signals generated by indirect electromagnetic wave paths in a conventional laboratory room. Flat samples of PTFE-4 were used as the measurement object.

**Findings:** New method for detecting the measured value against the background of interfering signals is proposed.

**Research limitations/implications:** Research on the possibility of using this method is ongoing, work is currently underway with low-directional antennas.

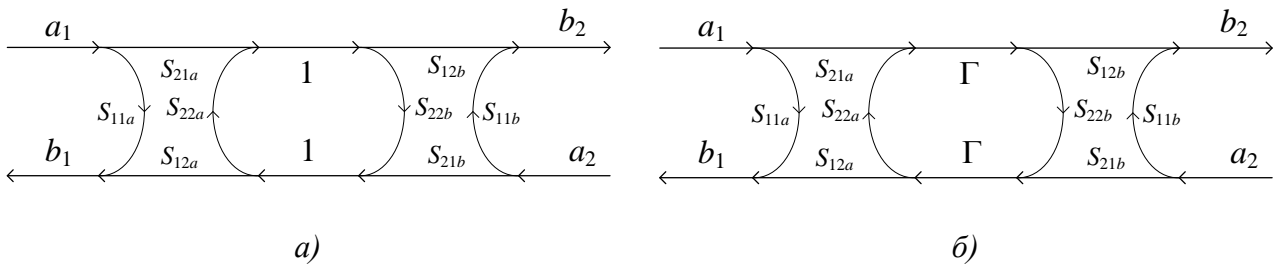
**Originality/value:** This method makes it possible to carry out antenna measurements without an anechoic chamber; in addition, processing the measurement results in a new way makes it possible to avoid the manifestation of the Gibbs effect.

**Keywords:** bistatic reflectance, highly directional antennas, measurements against the background of interfering signals



## Введение

Антенные измерения предполагают наличие дорогостоящего оборудования – прежде всего безэховой камеры для исключения влияния окружающего пространства на результаты измерений [1]. В работе [2] была предложена математическая модель бистатического коэффициента отражения электромагнитных волн от плоских образцов материалов в виде каскадного соединения 2-х портовых цепей, как показано на рис. 1. Калибровочная процедура была построена на знании коэффициентов отражения от приемной и передающей антенн со стороны их коаксиальных разъемов в режиме согласования с окружающим свободным пространством. Но при определении бистатического коэффициента отражения от контролируемых образцов материалов используется только измерение коэффициентов передачи, которое никак не связано с найденными коэффициентами отражения.



**Рис.1.Каскадное соединение цепей А и В:  
а – в направляющих системах, б – в открытом пространстве**

В данной работе предлагается иное решение на основе предложенной модели антенных измерений. Суть предлагаемого решения состоит не в определении параметров окружающих контролируемый объект цепей А и В, а в выявлении измеряемой величины на фоне мешающих сигналов, порожденных непрямыми трассами электромагнитных волн в обычном лабораторном помещении, без использования весьма дорогих поглощающих материалов клиновидной формы.

### Математическая модель измерения бистатического коэффициента отражения

Для измерения бистатического коэффициента отражения плоских образцов материалов и поглощающих покрытий необходимо выполнить измерения коэффициента передачи в фоновом (F) режиме, с металлическим отражателем (R) и с контролируемым объектом (X). Математическую модель таких измерений, в соответствии с моделью, представленной на рисунке 1, можно представить системой трех уравнений:

$$S_{21}^F = \frac{A\Gamma_F}{1 - B(\Gamma_F)^2}, S_{21}^R = \frac{A(\Gamma_F + \Gamma_R)}{1 - B(\Gamma_F + \Gamma_R)^2}, S_{21}^X = \frac{A(\Gamma_F + \Gamma_X)}{1 - B(\Gamma_F + \Gamma_X)^2}. \quad (1)$$

в которых  $S_{21}^F, S_{21}^R, S_{21}^X$  - результаты измерений коэффициентов передачи векторным анализатором цепей,  $\Gamma_F$  - коэффициент отражения, создаваемый фоновыми сигналами, приведенными в референсную плоскость,  $\Gamma_R$  - коэффициент отражения от плоского отражателя,  $\Gamma_X$  - коэффициент отражения от контролируемого объекта. Комплексные функции частоты электромагнитного поля А и В – моделируют влияние электромагнитных характеристик антенн и окружающего пространства на результаты измерений.

Решение системы уравнений (1) относительно искомой величины дает следующее соотношение:

$$\Gamma_X = \frac{\Gamma_R}{S_{21}^R} S_{21}^X \left[ \frac{\left(1 - B(\Gamma_F + \Gamma_X)^2\right)}{\left(1 - B(\Gamma_F + \Gamma_R)^2\right)} - S_{21}^F \frac{\left(1 - \frac{S_{21}^X}{S_{21}^R} \left(1 - B(\Gamma_F + \Gamma_X)^2\right)\right)}{\left(1 - \frac{S_{21}^X}{S_{21}^R} - B(\Gamma_R^2 + 2\Gamma_F \Gamma_R)\right)} \right] \quad (2)$$

Полученное соотношение привлекательно тем, что в нем отсутствует неизвестная величина А, а другая неизвестная В даже в случае остронаправленной, а тем более слабонаправленной антенны, во-первых, имеет весьма малое значение, а, во-вторых, является настолько быстроменяющейся функцией частоты, что всего лишь порождает шумоподобную дорожку на фоне медленно изменяющейся искомой величины  $\Gamma_X$  :

$$\Gamma_X(f) = \frac{\Gamma_R}{S_{21}^R(f)} \left( S_{21}^X(f) - S_{21}^F(f) \frac{\left(1 - S_{21}^X(f)/S_{21}^R(f)\right)}{\left(1 - S_{21}^F(f)/S_{21}^R(f)\right)} \right) \quad (3)$$

Здесь  $\Gamma_R$  - комплексный коэффициент отражения от плоского отражателя,  $\Gamma_R = -1$ , в случае вертикальной, и  $\Gamma_R = +1$ , случае горизонтальной поляризации зондирующей электромагнитной волны.

Чтобы убедиться в справедливости, полученного решения, были выполнены измерения контролируемых объектов с помощью остронаправленных антенн.

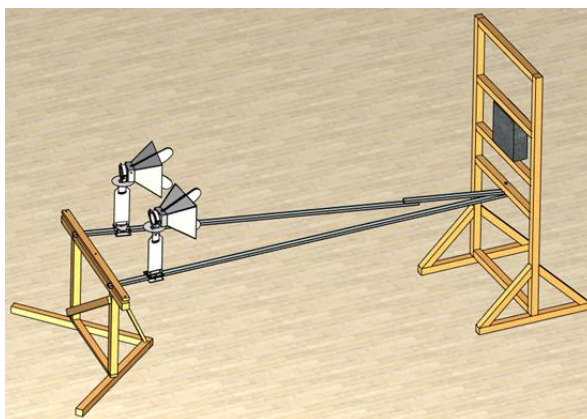
#### Антенный полигон для измерения бистатического коэффициента отражения

Схема антенного полигона показана на рис. 2. На основе измерения коэффициента передачи между передающей и приемной антеннами определяется коэффициент отражения от объекта измерения (бистатический коэффициент отражения). Расположенное за объектом наклонное зеркало предотвращает попадание сигналов, отраженных от объектов обстановки в приемную антенну. Для определения параметров стенда в открытом пространстве проводятся измерения коэффициента передачи с металлическим отражателем, установленным вместо объекта, и при отсутствии объекта.



Рис. 2. Схема измерительного стенда

Для обеспечения точного и стабильного позиционирования антенн разработана конструкция, показанная на рис. 3 и содержащая направляющие из профильных стальных труб сечением 20x40 мм, каретки для продольного смещения и стойки под разные типы антенн. Направляющие закреплены на специальных подставках, одна из которых служит также для размещения калибровочных мер и объектов измерения.



**Рис. 3. Конструкция измерительного стенда**

При экспериментах использованы узконаправленные зеркально-рупорные антенны, в качестве которых были использованы антенны спутникового телевидения с переделанным излучателем. Из него был изъят СВЧ приемник, а возбуждение волны осуществлялось штырем, введенным в круглый волновод.

Все измерения выполнялись в диапазоне частот от 10 до 17,8 ГГц анализатором цепей P4226/3 «Панорама», предварительно откалиброванным в коаксиальном тракте с использованием электронного калибратора

#### **Измерения плоских образцов материалов**

При использовании остронаправленных, например, параболических антенн падающая на плоский образец электромагнитная волна формируется в достаточно узком телесном угле, так что освещает лишь ограниченную площадь контролируемого плоского образца или зеркального отражателя. По этой причине дифракционные явления практически не проявляются и их влиянием на результаты контроля бистатического коэффициента отражения можно пренебречь.

Позиционирование антенн на первом шаге выполнялось с помощью рупорной антенны, как показано на рис.4. Рупорная антенна устанавливалась точно в центр предполагаемого пятна, после чего векторный анализатор цепей подключался к одной из параболических антенн и к рупору. По величине коэффициента передачи параболическая антенна фиксировалась в положении, где значение принятого рупорной антенной сигнала было максимальным.

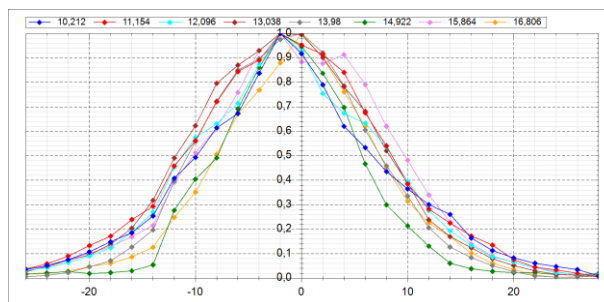


**Рис. 4. Позиционирование параболических антенн с помощью рупорной антенны**

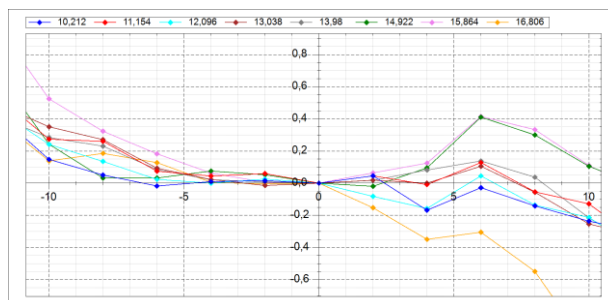
Окончательная юстировка антенн выполнялась по результатам полученных диаграмм распределения амплитуды и фазы зондирующей электромагнитной волны. Распределение амплитуды электромагнитного поля зондирующей электромагнитной волны в плоскости позиционирования объектов было получено с помощью измерения коэффициента отражения

от узкой металлической пластины, подвешенной вертикально и перемещаемой в горизонтальной плоскости, как показано на рисунке 5.

На рис. 5 приведены результаты измерения распределения по горизонтали амплитуды электромагнитного поля зондирующих электромагнитных волн в плоскости позиционирования контролируемых объектов для разных частот. Анализ графика показывает, что амплитуда имеет нормальный характер распределения на всех выбранных частотах. На рис.6 показано распределение фазы падающей волны по горизонтали в плоскости позиционирования контролируемых объектов для разных частот. На графике вертикальная ось – это радианы. Анализ графика показывает, что в процессе юстировки удалось добиться высокой точности позиционирования антенн и падающая на объект электромагнитная волна является квази-плоской.



**Рис. 5. Распределение амплитуды падающей волны для зеркальных антенн**



**Рис. 6. Распределение фазы падающей волны для зеркальных антенн**

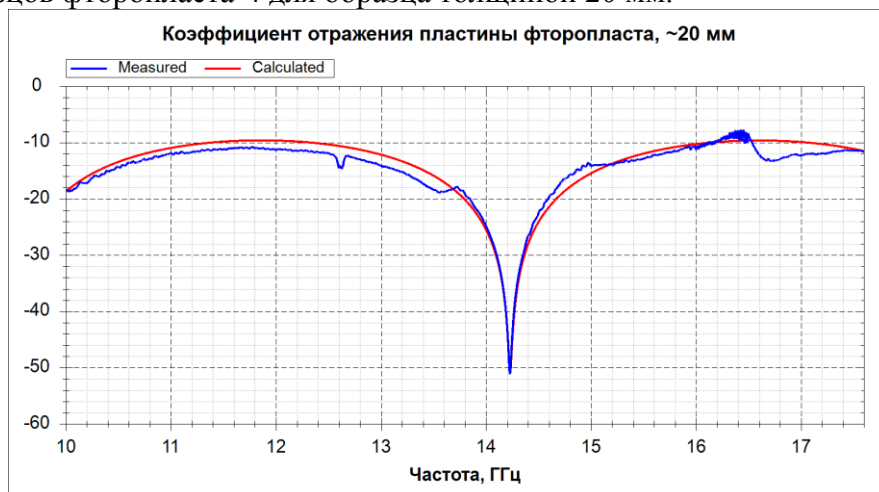
Были измерены бистатические коэффициенты отражения плоских образцов фторопласта-4 и оргстекла разной толщины. Ниже приведено соотношение для бистатического коэффициента отражения, составленное в соответствии с известными формулами Френеля [3] при наклонном падении плоской волны на границу пластины толщиной  $d$  с комплексной диэлектрической проницаемостью  $\hat{\epsilon} = \epsilon' - i\epsilon''$  при вертикальной поляризации электрического поля:

$$S_{11} = \frac{\cos \phi - \sqrt{\hat{\epsilon}} \cos \vartheta}{\cos \phi + \sqrt{\hat{\epsilon}} \cos \vartheta} - \frac{4 \cos \phi \cos \vartheta \sqrt{\hat{\epsilon}}}{\left( \cos \phi + \sqrt{\hat{\epsilon}} \cos \vartheta \right)^2} \frac{\cos \phi - \sqrt{\hat{\epsilon}} \cos \vartheta}{\cos \phi + \sqrt{\hat{\epsilon}} \cos \vartheta} \exp\left(-i4\pi f d \sqrt{\hat{\epsilon}} / (v_0 \cos \vartheta)\right) \quad (4)$$

$$1 - \left( \frac{\cos \phi - \sqrt{\hat{\epsilon}} \cos \vartheta}{\cos \phi + \sqrt{\hat{\epsilon}} \cos \vartheta} \right)^2 \exp\left(-i4\pi f d \sqrt{\hat{\epsilon}} / (v_0 \cos \vartheta)\right)$$

Здесь  $\phi = 8,085^{\circ}$  - угол падения,  $\vartheta$  - угол преломления, который определяется вторым законом Снеллиуса:  $\sin \vartheta / \sin \phi = 1 / \sqrt{\epsilon'}$ .

На рис. 7 приведены экспериментальные и теоретические значения коэффициента отражения образцов фторопласта-4 для образца толщиной 20 мм.



**Рис. 7. Измеренный и рассчитанный коэффициент отражения пластины фторопласта-4 толщиной 20 мм**

В табл. 1 приведены полученные значения действительной и мнимой относительной диэлектрической проницаемости для разных толщин образцов. Из таблицы видно, что для образцов из фторопласта-4 значения эффективной диэлектрической проницаемости для образцов разных толщин практически совпадают, как и толщины образцов.

Образцы оргстекла были составлены из отдельных пластин толщиной 4 и 6 мм, между которыми образовывались воздушные зазоры, этим можно объяснить, почему рассчитанная толщина образцов оказалась ниже ожидаемой.

**Таблица 1.**  
**Сравнение параметров образцов из оргстекла и фторопласта с теоретическими данными**

Образец, мм	$\dot{\epsilon}$					Толщина, мм		
	$\epsilon'_{изм}$	$\epsilon''$	$tg\delta_{изм}$	$\epsilon'_{теор}$	$tg\delta_{теор}$ при 1МГц	d	d <sub>min</sub>	d <sub>max</sub>
<b>фторопласт-4</b>								
~6	1,995	0,0038	0,0019	1,9-2,1	(2-2,5) * 10 <sup>-4</sup>	5,48	5,34	5,83
~10	1,990	0,0035	0,00175			10,75	10,68	11,4
~20	1,992	0,0035	0,00175			22,480	21,9	22,66
<b>оргстекло</b>								
~6	3,020	0,21137	0,069	3,5-4,2	0,02-0,06	5,329	6	3,020
~12	3,240	0,07374	0,023			10,168	12	3,240
~16	3,134	0,0478	0,015			14,039	16	3,134
~20	2,933	0,01899	0,006			18,041	20	2,933

### Заключение

Предложен новый метод измерения бистатического коэффициента отражения плоских образцов и поглощающих покрытий на фоне помех окружающего пространства. Метод позволяет практически полностью исключить влияние фонового сигнала из результатов измерений. Метод апробирован на узконаправленных антеннах. Результаты, полученные для хорошо известных материалов, таких как фторопласт-4 и оргстекло, совпадают с теоретическими данными, что подтверждает работоспособность предложенного метода.

### Библиографический список

1. **Меньших, Н.Л.** Уменьшение ошибки позиционирования при измерении бистатического модуля коэффициента отражения / Меньших Н.Л., Федоров С.А. // VII Всероссийская микроволновая конференция – 2020 - 2020., С. 363-366
2. **Евсеев, В.И.** Методы и средства контроля параметров объектов в нестандартных волноведущих системах и в открытом пространстве. Диссертация на соискание ученой степени кандидата технических наук. - Нижний Новгород, НГТУ им. Р.Е. Алексева. - 2021
3. **Никольский, В.В.** Электродинамика и распространение радиоволн. М.: Наука: Физматлит, 1978.

## МЕТОДИКА И АЛГОРИТМ МОНИТОРИНГА И АДАПТАЦИИ СТРУКТУРЫ СИСТЕМЫ СВЯЗИ ОРГАНИЗАЦИИ

С.С. Белоусов

ORCID: 0000-0002-4314-6806 e-mail: belousovtel@yandex.ru

П.А. Федюнин

ORCID: 0000-0001-6663-4362 e-mail: fpa1969@yandex.ru

Военный учебно-научный центр Военно-воздушных сил «Военно-воздушная академия им. проф. Н. Е. Жуковского и Ю. А. Гагарина» (г. Воронеж)

В работе предложена методика и алгоритм мониторинга и адаптации структуры системы связи организации, основанные на определении функционального состояния направлений связи между подчиненными объектами в режиме реального времени и гибком изменении указанной структуры в соответствии со складывающейся обстановкой. Проведен количественный и качественный анализ среднего времени исправности требуемого ресурса связи на каждом пункте управления входящим в единую сеть управления объектами, путем расчёта живучести, надежности и помехоустойчивости отдельных линий и каналов связи. Кроме того определены меры по восстановлению исправности системы связи за счёт проведения мероприятий по вводу резерва, срочных ремонтных мероприятий, а также ввода мобильных комплектов аппаратуры.

**Ключевые слова:** мониторинг, адаптация, управление, структура, система связи, исправность, восстановление.

## METHODOLOGY AND ALGORITHM FOR MONITORING AND ADAPTATION OF THE STRUCTURE OF THE COMMUNICATION SYSTEM OF ORGANIZATION

S.S. Belousov, P.A. Fedyunin

Zhukovsky and Gagarin Voronezh Air Force Academy

**Purpose:** The author proposes a methodology and algorithm for monitoring and adapting the structure of an organization's communication system, based on determining the functional state of communication directions between subordinate objects in real time and flexible changing this structure in accordance with the current situation.

**Design / methodology / approach:** The article selects the health indicators of the technical basis of the organization's management system - the communication system, as well as the classification of functional states, which makes it possible to determine the most optimal object control point at the current time, which has the best characteristics. The author in his methodology applies the provisions of the theory of reliability, the theory of probability, determining the functional states of the communication system of an organization based on a graph with the compilation of Kolmogorov's differential equations. This solution allows you to determine and timely adapt the structure of the communication system in order to improve the efficiency of managing subordinate objects.

**Findings:** The author suggests promising directions for the development of research to improve the efficiency of facility management, taking into account the obtained scientific results.

**Research limitations/implications:** The study is limited to a predictable database of destructive external influences on the control system, which has a certain random character.

**Originality/value:** The value of the entire article is the novelty in the approach to determining the functionality of a communication system based on the serviceability of its elements, as well as its timely and flexible rebuilding, which allows to ensure the continuity of control of subordinate objects in the interests of performing their tasks.

**Key words:** monitoring, adaptation, management, structure, communication system, serviceability, restoration.

## Введение

Выполнение поставленных задач силами организации 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 [1].

Под непрерывностью управления организацией в статье понимается способность системы управления обеспечить процесс управления в необходимом темпе и функционирование управляемых сил без снижения их готовности, устойчивости и эффективности [2].

Существующее состояние системы управления организацией в ходе выполнения задач по предназначению характеризуется значительными изменениями ее технической основы – системы связи, связанными с деструктивным воздействием противоборствующей стороны, приводящим в определенных условиях к потере (дезорганизации) управления. Целью статьи является разработка модели мониторинга и адаптации функционального состояния системы связи в целях обеспечения непрерывности управления организацией. Определение зависимости вероятности непрерывности управления от интенсивности потери (восстановления) управления.

### Влияние состояния системы связи на процесс управления организацией

Как известно, любая система включает совокупность различных элементов, находящихся в отношениях и связях между собой и образующих определенную целостность. Это положение в полной мере относится и к системе управления организацией [3]. Технической основой указанной системы является система связи (рис. 1), которая выполняет функции по обеспечению обмена данными между элементами системы управления и степени реализации ее возможностей, во многом, определяется результат, который может быть достигнут организацией.



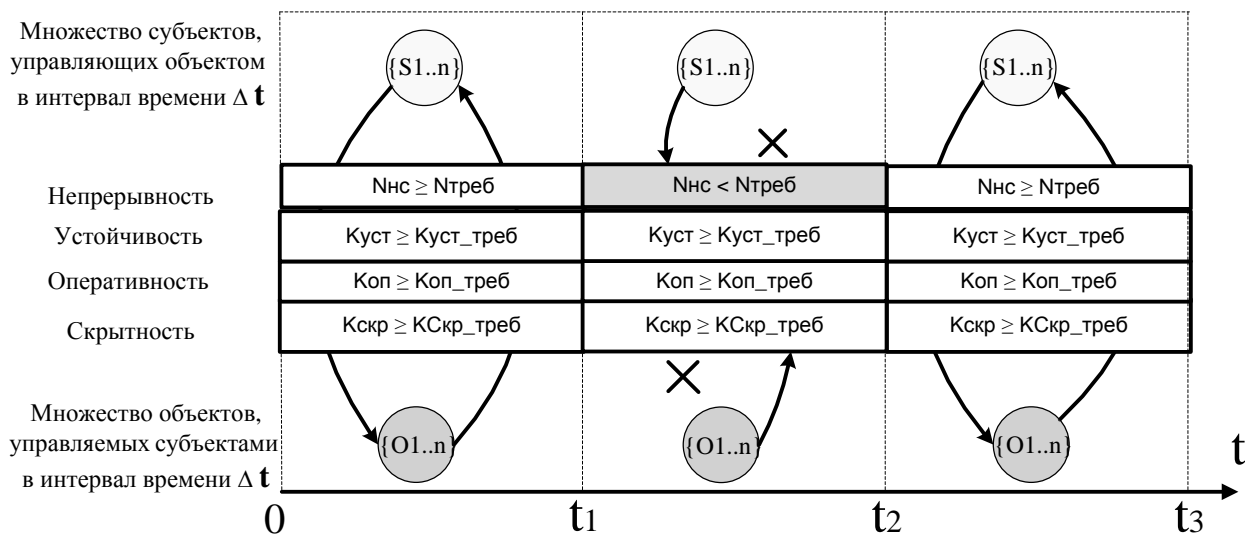
Рис. 1. Структура системы управления организацией

Наибольшее влияние на непрерывность управления организацией оказывают устойчивость и пропускная способность системы связи, ввиду зависимости указанного качества управления от временных интервалов потери управления и времени на его восстановление [4].

Для обеспечения потребности системы управления в информационном обмене и решения задач управления на базе системы связи создаются информационные направления, которые реализуются направлениями связи. Следовательно, 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 она создается в целях решения задач управления). Поэтому, в целях оценки влияния системы связи на качество управления наиболее значимыми в модели используются показатели, характеризующие количество и качество функционирующих направлений связи.

На рис. 2 показано влияние состояния системы связи на процесс управления на условном интервале времени выполнения задачи от момента  $t_0$  до  $t_3$ . Управление будет непрерывным, если система связи обеспечит функционирование требуемого числа направлений связи  $N_{\text{нс}} \geq N_{\text{треб}}$ , указанное требование выполняется на интервалах с  $t_0$  до  $t_1$ , а также с  $t_2$  до  $t_3$ . В период с  $t_1$  до  $t_2$  в результате невыполнения требований по функционированию заданного количества направлений связи  $N_{\text{нс}} < N_{\text{треб}}$  управление не обеспечено, и восстановлено только с момента  $t_2$ .





**Рис. 2. Влияние состояния системы связи на процесс управления организацией**

Непрерывность управления будет обеспечена, если система связи будет способна своевременно гибко реагировать на изменения обстановки (деструктивное воздействие противника) при наличии требуемого ресурса связи на всем временном интервале выполнения поставленной задачи с  $t_0$  до  $t_3$  [5].

Как известно, построение системы связи определяется организацией управления, определением мест развертывания пунктов управления, в том числе по этапам выполнения поставленной задачи, а также порядок передачи управления с одного пункта управления на другой при выходе из строя элементов системы управления. В целях анализа смены структуры системы связи во времени выделим возможные её состояния:

1. Требуемое (планируемое) состояние системы связи, организуемое на основании положений руководящих документов по вопросам обеспечения управления, гарантированно обеспечивающее условие  $N_{нс} > N_{треб}$ .

2. Допустимое состояние системы связи, организуемое на основании схемы узлу связи, с учетом текущей обеспеченности и состояния средств связи рассматриваемого субъекта управления и позволяющая еще по количеству функционирующих направлений связи обеспечить непрерывность управления организацией, т.е.  $N_{нс} = N_{треб}$ .

3. Прогнозируемое состояние системы связи, представляющее собой ожидаемое состояние ресурса системы связи на рассматриваемых элементах системы управления, с учетом вероятного воздействия и влияния других различных внешних и внутренних факторов обстановки. Полагая при этом, что высока вероятность наличия моментов, когда  $N_{нс} < N_{треб}$ .

В условиях внешнего воздействия изменение структуры системы пунктов управления прогнозируемо влечет за собой изменение технической основы системы управления – системы связи. Однако, этот процесс происходит с определенной степенью инерции, связанной с затратами времени на переключение каналов связи, включение резервной аппаратуры связи, согласования вопросов организации связи между узлами, а также наличием внешних и внутренних факторов, влияющих на непрерывность управления организации. Также необходимо учитывать, что при перестроении системы управления и передаче управления на другой пункт управления, возможности запасных пунктов управления по управлению объектами, как правило, меньше чем у основного пункта управления. Данное обстоятельство еще более усложняет гарантированное выполнение требования к системе связи по опережающей готовности технической основы системы управления относительно органа управления.

Учитывая указанные обстоятельства, приоритетным направлением в интересах недопущения потери управления на  $j$ -м ПУ является установление момента времени, когда система связи не обеспечит наличие у органа управления требуемого ресурса ( $N_{нс} \geq N_{треб}$ ). При

этом, система управления должна своевременно осуществить контроль состояния ресурса связи (направлений связи) на тех элементах системы управления, с которых согласно установленного порядка осуществляться решение задач управления, а также давать рекомендации по смене пункта управления либо по изменению состояния системы связи в интересах реализации установленного порядка управления в целях обеспечения его непрерывности.

### **Модель мониторинга и адаптации функционального состояния системы связи**

В целях решения задачи по обеспечению непрерывности управления авиационным формированием за счет обеспечения функционирования необходимого числа направлений связи, а также предоставления время для восстановления состояния нарушенной структуры системы связи авторами предлагается структурная схема модели мониторинга и адаптации функционального состояния системы связи [6] представленная на рис. 3.



**Рис. 3. Структурная схема модели мониторинга и адаптации функционального состояния системы связи организации (вариант)**

Для обеспечения непрерывности управления с пункта управления система связи должна иметь высокие значения показателей, характеризующих ее готовность, мобильность, устойчивость, пропускную способность и разведывательную защищенность.

Из всех перечисленных свойств системы связи предлагается рассмотреть устойчивость и пропускную способность, оказывающих наибольшее влияние на непрерывность управления. Поддержание заданных уровней устойчивости и пропускной способности системы связи является необходимым условием для обеспечения непрерывности управления. Наибольшее влияние на непрерывность управления организацией оказывают устойчивость и пропускная способность системы связи, ввиду зависимости указанного качества управления от временных интервалов потери управления и времени на его восстановление [4].

Устойчивость системы связи характеризует ее способность обеспечивать управление силами ИА в условиях воздействия на ее элементы различных видов оружия противника, опасных факторов техногенного и природного характера и помех всех видов.

Устойчивость системы связи определяется ее живучестью, помехоустойчивостью и надежностью. Поскольку устойчивость системы связи определяется совокупностью воздействия ряда факторов, необходимо учитывать степень каждого из них в различных условиях обстановки. Действия факторов, обуславливающих живучесть, помехоустойчивость и надежность системы связи или ее элементов, можно с достаточной степенью точности считать независимыми. Тогда:

$$K_u^{лс} = K_{иэс}^{лс} K_{ипу}^{лс} K_{ин}^{лс}, \quad (1)$$

где  $K_{иэс}^{лс}, K_{ипу}^{лс}, K_{ин}^{лс}$  - коэффициенты исправного действия системы (элемента системы) связи по живучести, помехоустойчивости, надежности.

В целом высокая устойчивость системы связи достигается выполнением всего комплекса мероприятий, направленных на повышение живучести, помехоустойчивости и надежности системы связи. В свою очередь, пропускная способность системы связи характеризует потенциальные возможности ее по передаче потоков сообщений в единицу времени. Количественно пропускная способность направления связи в статье оценивается интенсивностью входящего потока сообщений с заданной своевременностью. Полученные значения указанных показателей позволяют оценить возможности системы связи по обеспечению непрерывности управления организацией.

### Выводы

Таким образом, предлагаемая модель мониторинга и адаптации функционального состояния системы связи позволит органу управления с учетом анализа сложившейся обстановки (наличия на предлагаемых пунктах управления оперативного состава, сил непосредственно выполняющих поставленную задачу, обеспечивающих сил, а также других условий и факторов) повысить оперативность выбора наиболее рационального по состоянию ресурса связи пунктов управления и тем самым обеспечить качество управления организацией.

### Библиографический список

1. Белоусов, С. С. Математическая модель оценки возможностей системы связи по обеспечению непрерывности управления авиационным формированием / С. С Белоусов, П. А. Федюнин // сб. мат. XXI Международной конференции «Информатика: проблемы, методы, технологии» (IPMT-2021) (Воронеж, 11–12 февраля 2021 г.). – Воронеж, 2021. – С. 1696-1698.
2. Алтухов, П. К. Основы теории управления войсками/ И. А. Афонский, И. В. Рыболовский, А. Е. Татарченко; Под ред. П. К. Алтухова. – М. : Воениздат, 1984. – 221 с.
3. Боговик, А. В., Теория управления в системах военного назначения: учебник./ А.В. Боговик, В.В. Игнатов – СПб.: ВАС. 2008. – 460 с.
4. Белоусов, С. С. Подход к построению пространственно-временной модели управления подразделением истребительной авиации при выполнении задач по предназначению / С. С Белоусов, Ю. Е. Донсков, П. А. Федюнин // Военная Мысль. - 2021. - № 5. - С. 65-76.
5. Белоусов, С. С. Проведение исследований непрерывности управления авиационным формированием с использованием средств компьютерного моделирования / С. С Белоусов, П. А. Федюнин, С. П. Юрченко // сб. мат. XXI Международной конференции «Информатика: проблемы, методы, технологии» (IPMT-2021) (Воронеж, 11–12 февраля 2021 г.). – Воронеж, 2021. – С. 1623-1627.
6. Свидетельство о государственной регистрации программы для ЭВМ № 2021661142 от 06.07.2021 г Программа для проведения мониторинга и адаптации функционального состояния системы связи. / Белоусов С.С.; заявитель и правообладатель Белоусов С. С., заявка №. 2021619702; в реестре программ для ЭВМ 06.07.21.

## МОДЕЛЬ И АЛГОРИТМ ФУНКЦИОНИРОВАНИЯ АВТОМАТИЗИРОВАННОЙ СЕТИ ДЕКАМЕТРОВОЙ РАДИОСВЯЗИ С АДАПТИВНЫМ УПРАВЛЕНИЕМ СТРУКТУРНЫМ И ПАРАМЕТРИЧЕСКИМ РЕСУРСОМ

А. В. Рябов

ORCID: 0000-0003-4578-0840 e-mail: [ryalvy@mail.ru](mailto:ryalvy@mail.ru)

А. Е. Алексеев

e-mail: [sanei@bk.ru](mailto:sanei@bk.ru)

Военный учебно-научный центр Военно-воздушных сил «Военно-воздушная академия им. проф. Н. Е. Жуковского и Ю. А. Гагарина» (г. Воронеж)

Представлены модель и алгоритм функционирования автоматизированной сети декаметровой радиосвязи с адаптивным управлением структурным и параметрическим ресурсом. Структурная адаптация предлагается на основе ретрансляции сигналов по составным или параллельным радиолиниям. Управление структурой радиолинии в предложенной модели осуществляется на основе исходных данных о состоянии ионосферы, сигнально-помеховой обстановки в точке приема, топологии сети и состоянии ее элементов. В рамках структурной адаптации обеспечивается возможность управления параметрами радиолинии с целью их оптимизации. Разработанный алгоритм функционирования предложенной радиосети позволяет на основе анализа результатов исходных данных обеспечить работу радиосети в режиме организации радионаправлений в простых условиях сигнально-помеховой обстановки, либо в одном из адаптивных помехоустойчивых режимов. В модели предложены следующие режимы повышения помехоустойчивости: ретрансляция сигналов по составным либо параллельным радиолиниям с выбором оптимального количества ретрансляторов и оптимизацией их энергетических параметров для обеспечения требуемой достоверностью связи; попеременная ретрансляция сигналов с минимизацией времени излучения передатчиков в целях повышения разведзащищенности радиосети; комбинированные режимы, основанные на сочетании перечисленных. Разработанные модель и алгоритм функционирования автоматизированной сети декаметровой радиосвязи позволяют обосновать способы повышения помехоустойчивости и разведзащищенности на основе адаптивного управления структурным и параметрическим ресурсом.

**Ключевые слова:** декаметровая радиосвязь, автоматизированная сеть радиосвязи, адаптивное управление, структурная адаптация, параллельная ретрансляция, попеременная ретрансляция, помехоустойчивость, разведзащищенность

## MODEL AND ALGORITHM OF FUNCTIONING OF AN AUTOMATED NETWORK OF DECA-METAL RADIO COMMUNICATION WITH ADAPTIVE CONTROL OF STRUCTURAL AND PARAMETER RESOURCE

A. V. Ryabov

Zhukovsky and Gagarin Voronezh Air Force Academy

**Purpose:** Improving the noise immunity and intelligence protection of decameter radio communication networks based on adaptive control of the structural and parametric resource.

**Design/methodology/approach:** a model and an algorithm for the operation of an automated decameter radio communication network with adaptive control of a structural and parametric resource have been developed. Structural adaptation is proposed on the basis of signal retransmission over composite or parallel radio links. The control of the radio link structure in the proposed model is based on the initial data on the state of the ionosphere, the signal-interference situation at the receiving point, the network topology and the state of its elements. Within the framework of structural adaptation, it is possible to control the parameters of the radio link in order to optimize them.

**Findings:** The presented model and the functioning algorithm of the adaptive radio network make it possible to estimate the possibilities for increasing its effectiveness.

**Research limitations/implications:** the presented article is the starting point for further research in the field of building adaptive radio networks.

**Originality/value:** the developed model and algorithm for the operation of an automated decameter radio communication network make it possible to justify ways to increase noise immunity and intelligence protection based on adaptive control of a structural and parametric resource.

**Keywords:** decameter radio communication, automated radio communication network, adaptive control, structural adaptation, parallel relaying, variable relaying, noise immunity, intelligence protection

## Ведение

В настоящее время построение автоматизированных радиосетей с автоматическим установлением и адаптивным ведением ДКМ радиосвязи основывается, как правило, на использовании следующих стандартов информационного обмена: MIL-STD-188-141, HFDL, STANAG 4539, STANAG 5066 [1-4]. Перечисленные стандарты регламентируют процедуры автоматического составления каналов, применение сигнально-кодовых конструкций и их адаптацию к состоянию канала в процессе сеанса связи, а также сетевые проблемы многопользовательского использования ДКМ радиоканалов [1-4].

Реализация рассматриваемых стандартов требует решения следующих задач [1-4]:

назначение набора частот и автоматическое составление канала связи с автоматическим выбором рабочей частоты;

использованием адаптивных методов приема для передачи информации с высокими скоростями;

повышение достоверности связи на основе помехоустойчивого кодирования с применением исправляющих кодов, а также на основе применения канала обратной связи для автоматического запроса повторения искаженных пакетов данных;

скремблирование и перемежение данных для обеспечения имитозащиты и борьбы с пакетированием ошибок из-за замираний и импульсных помех;

многопараметрическая адаптация по частоте, скорости передачи информации, мощности, виду модуляции, способу кодирования.

При этом рассматриваемые стандарты предусматривают постоянный автоматический мониторинг ионосферы для оперативного определения оптимальной рабочей частоты и позволяют достичь надежности в канале не менее 0,90–0,95 со скоростями передачи данных от 75 бит/с до 9,6 Кбит/с [1]. Однако протоколы этих стандартов позволяют установить соединение по типу радионаправления или, в частном случае, организовать радиальную сеть с выделением главной станции. Недостатком таких протоколов является невозможность организации ДКМ радиосетей по полносвязной топологии.

Топология полносвязной сети позволяет реализовать методы адаптивного управления ее структурой на основе применения ретрансляции сигналов по составным или параллельным радиолиниям [5, 6]. Реализация такой радиосети предлагается на основе объединения отдельных средств ДКМ радиосвязи, распределенных территориально в единую полносвязную сеть.

### **Модель автоматизированной сети декаметровой радиосвязи с адаптивным управлением структурным и параметрическим ресурсом**

Объединение средств ДКМ радиосвязи в адаптивную автоматизированную радиосеть обеспечивает возможность применения методов структурной адаптации, например ретрансляции через составные радиолинии, а также параллельной ретрансляции с синфазным сложением мощности сигналов в точке приема и синхронизацией времени излучения ретрансляторов по каналу обратной связи [5]. При этом предлагаемой сети ДКМ радиосвязи предполагается работа каждой радиостанции в следующих режимах [5, 6]:

режим организации прямых связей между взаимодействующими пунктами управления;

ретрансляция сигналов по составным радиолиниям в пределах сети;

одновременная (параллельная) ретрансляция сигналов двумя и более радиостанциями с синфазным сложением мощности в точке приема и синхронизацией передатчиков ретрансляторов по каналу обратной связи;

ретрансляция сигналов двумя и более радиостанциями с попеременным (поочередным) излучением сигналов.

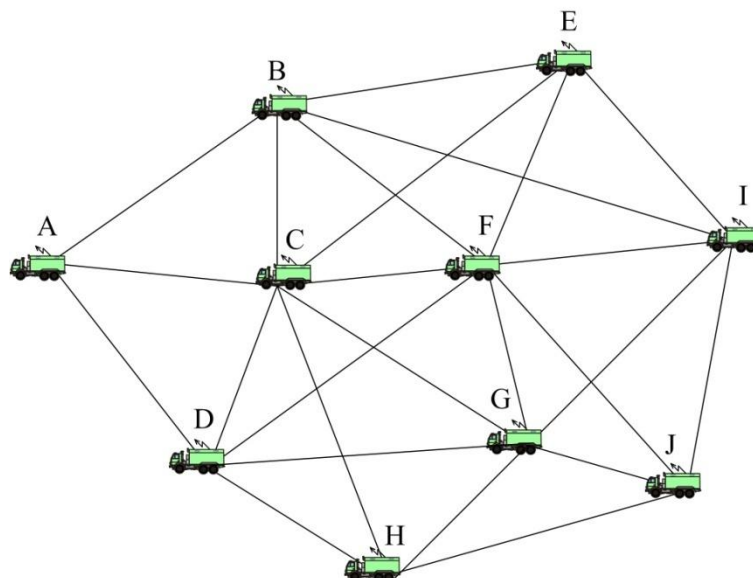
В простой сигнально-помеховой обстановке, работа рассматриваемой адаптивной сети радиосвязи может осуществляться, в режиме организации прямых связей между взаимодействующими пунктами управления, либо режиме последовательной ретрансляции сигналов через составные радиолинии. При этом в сети необходим постоянный мониторинг доступного структурного ресурса и оценка радиочастотной обстановки.

При ухудшении сигнально-помеховой обстановки, при наличии доступных ретрансляторов, возможен переход в режим параллельной ретрансляции. В этом случае осуществляется прием сигналов по параллельным каналам с синфазным сложением мощности и по каналу обратной связи осуществляется синхронизация передатчиков ретрансляторов. Также, для снижения

электромагнитной доступности излучающих радиосредств возможен переход попеременной ретрансляции сигналов двумя и более радиостанциями [5, 6].

Для построения автоматизированной полносвязной автоматизированной сети ДКМ радиосвязи, реализующей многоуровневую адаптацию и вариативную маршрутизацию, требуется разработка протокола маршрутизации. Рассмотрим принцип функционирования такого протокола.

Перед началом работы радиосети на всех станциях размещается информация о ее топологии – точный граф связей сети, в котором вершинами графа являются ДКМ радиостанции, а ребрами – радионаправления между ними (рисунок 1).



**Рис. 1 – Вариант структуры автоматизированной сети ДКМ радиосвязи**

В протоколе маршрутизации определяются взаимодействующие станции. В дальнейшем взаимодействующие станции обмениваются с информацией о текущем графе связей сети. При получении информации о графе связей от взаимодействующей станции, она сохраняется и рассылается другим станциям, за исключением той, от которой была получена. Это позволит избежать закливания устаревшей информации о состоянии графа связей сети между взаимодействующими станциями. В результате все станции сохраняют в своей памяти идентичные сведения о текущей конфигурации графа связей сети.

Для контроля состояния связей и работоспособности сети, взаимодействующие станции обмениваются короткими формализованными сообщениями через установленный временной интервал. При прекращении поступления контрольных сообщений от какой-либо станции в течение нескольких заданных временных интервалов, делается вывод об изменении состояния связи на неработоспособное и вносится изменение в топологическую базу данных. Одновременно взаимодействующим станциям передается обновленная информация о графе состояния сети, для уточнения баз данных.

Основной задачей протокола маршрутизации является составление маршрута между двумя станциями. Для ее решения обычно используется итеративный алгоритм Дейкстры [7]. В соответствии с этим алгоритмом, каждая станция определяет оптимальные маршруты до всех станций сети. При изменении информации о графе связей сети, каждая станция пересчитывает оптимальные маршруты и корректирует свою таблицу маршрутизации.

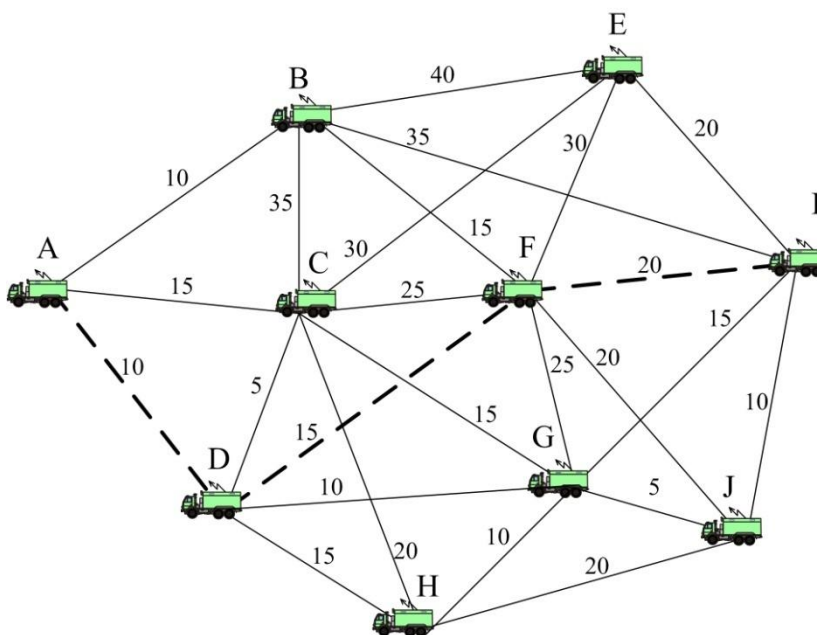
Установление связи между двумя станциями сети по протоколу маршрутизации начинается с выбора метрики. По умолчанию используется метрика, учитывающая пропускную способность радионаправлений. Кроме того, допускается применение метрик, учитывающих надежность и качество приема, которое зависит от мощности помех на входе приемников.

Для каждой из метрик протокол строит отдельную таблицу маршрутизации. Перед началом работы информация о графе связей сети загружается в маршрутизатор станции. Оператором станции-отправителя выбирается метрика, например качество связи. Каждому ребру графа связей сети назначаются количественные значения. Пример назначения количественных значений представлен в таблице 1.

Вариант количественных значений графа связей сети

Отношение сигнал/шум (дБ)	10	20	30	40	50	60	70	80	90	100
Значение метрики	50	45	40	35	30	20	15	10	5	1

Исходя из назначенной метрики, в соответствии с алгоритмом Дейкстры, протокол определяет кратчайший маршрут (с наименьшей метрикой). На рисунке 2 представлен возможный вариант работы протокола маршрутизации автоматизированной сети ДКМ радиосвязи. Кратчайший маршрут обозначен пунктирной линией (A, D, F, I).



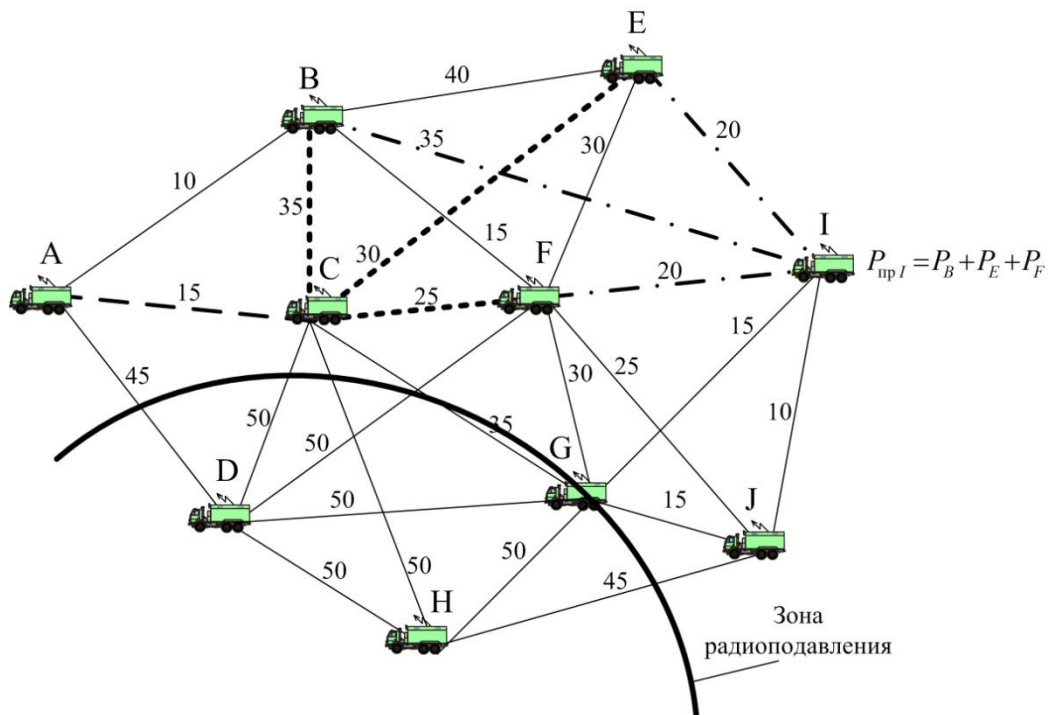
**Рис. 2. Вариант функционирования протокола маршрутизации автоматизированной сети ДКМ радиосвязи**

В случае снижения качества связи на каком-либо участке маршрута, либо выхода из строя какой-либо станции сети, в соответствии с алгоритмом функционирования сети, маршрут передачи пересчитывается. При этом для оптимизации показателей помехоустойчивости часть радиостанций сети может быть переведена в режим параллельной ретрансляции сигналов. В этом случае назначаются станции отправителя и получателя, производится поиск свободных ретрансляторов и определение их необходимого количества [6].

Описываемая структура представляет собой статическую маршрутизацию для заданного набора станций. Вариант функционирования протокола маршрутизации кластера ДКМ радиосвязи с применением параллельной ретрансляции сигналов представлен на рисунке 3.

В варианте, представленном на рисунке 3, в зоне радиоподавления находятся станции D, H, G. При очередной оценке состояния связей и работоспособности соседних станций это отразится либо отсутствием контрольных сигналов от этих станций при их подавлении, либо в резком снижении качества связи с ними. В этом случае взаимодействующие станции вносят изменения в свои базы данных и рассылают информация о графе связей остальным станциям сети.





**Рис. 3. Вариант функционирования протокола маршрутизации сети ДКМ радиосвязи с применением режима параллельной ретрансляции**

Оператор переводит автоматизированную сеть в режим параллельной ретрансляции, с назначением станции отправителя (С), станции получателя (I) и станций ретрансляторов (В, Е, F). При этом, мощность излучения передатчиков назначенных ретрансляторов В, Е и F снижается пропорционально их количеству (для рассматриваемого примера – в 3 раза), что повысит их разведзащищенность. Конечный вариант маршрута для рассматриваемого примера выделен пунктирной линией (на рисунке 3 – А, С, {В, Е, F}, I).

### **Алгоритм функционирования автоматизированной сети декаметровй радиосвязи с адаптивным управлением структурным и параметрическим ресурсом**

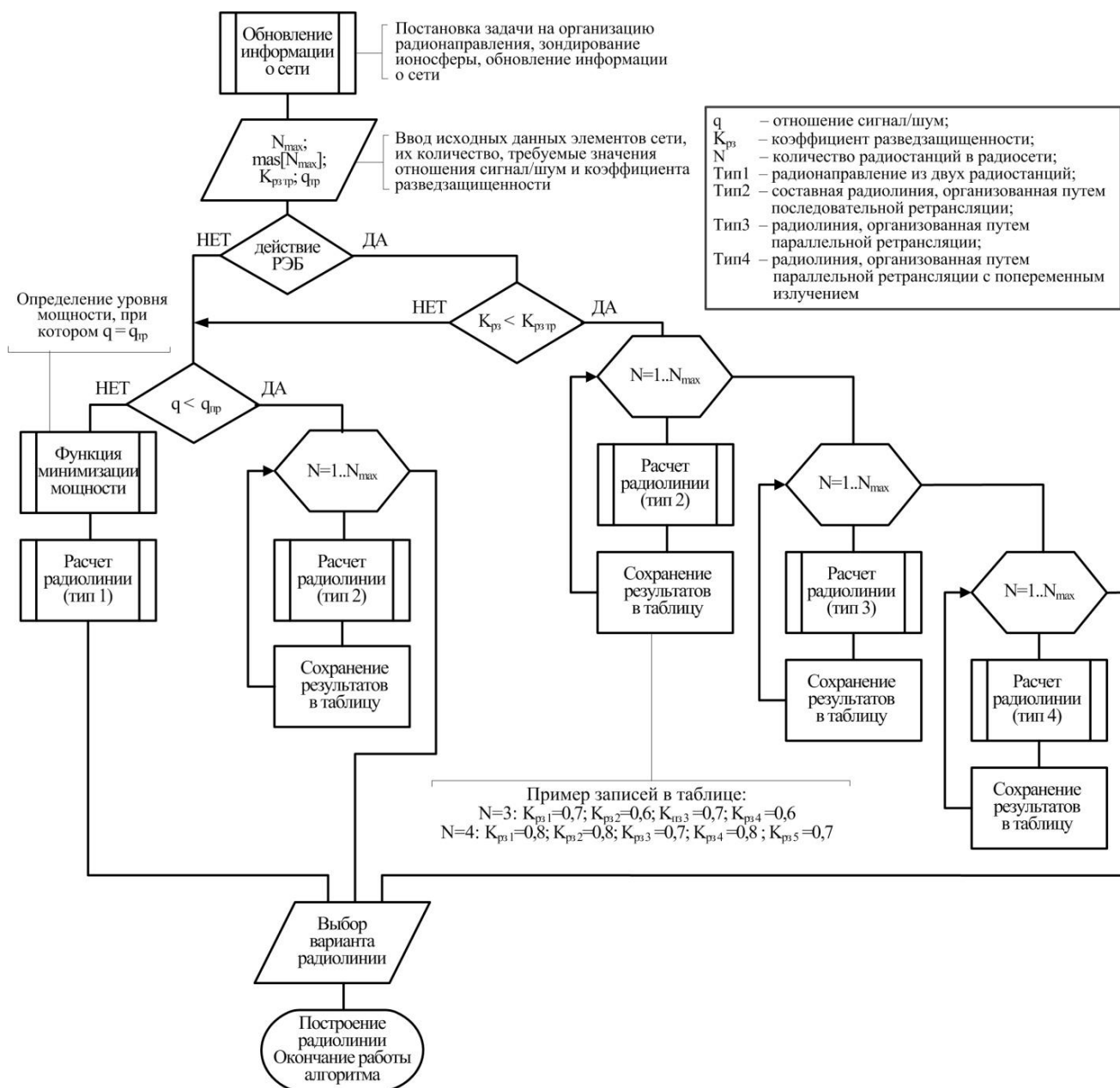
Рассмотрим алгоритм функционирования автоматизированной сети ДКМ радиосвязи с адаптивным управлением структурно-параметрическим ресурсом, представленной на рисунках 1-3. Данная радиосеть представляет собой совокупность средств ДКМ радиосвязи, распределенных на определенной территории.

В общем случае каждый элемент такой 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 через составные радиолинии информации от источника получателю;
- параллельная ретрансляция информации от источника получателю с синфазным сложением мощности излучения передатчиков в точке приема [5, 6];
- попеременная ретрансляция сигналов информации от источника получателю системой пространственно распределенных ретрансляторов (радиостанций) [5, 6].

На рисунке 4 представлен алгоритм поиска оптимального маршрута прохождения радиосигналов в сети ДКМ радиосвязи, с учетом перечисленных вариантов применения.





**Рис. 4. Алгоритм поиска оптимального маршрута прохождения радиосигналов в автоматизированной сети ДКМ радиосвязи с адаптивным управлением структурно-параметрическим ресурсом**

Отличиями предлагаемого алгоритма от существующих являются:

возможность проведения расчета показателей разведзащищенности по методике, рассмотренной в [7,8], как для радионаправления типа «точка-точка», так и для радиосети любой топологии, в том числе с адаптивным управлением структурно-параметрическим ресурсом;

возможность оценки эффективности применения способов повышения разведзащищенности системы связи на основе адаптивного управления ее структурно-параметрическим ресурсом, с выбором оптимального способа повышения разведзащищенности и маршрута прохождения сигнала;

возможность оценки качества радиосвязи (через соотношение сигнал/шум) в простой сигнально-помеховой обстановке, с учетом оптимизации энергетических и временных показателей средств радиосвязи для поддержания требуемой разведзащищенности сети.

На рисунке 5 представлена используемая в приведенном на рисунке 4 алгоритме функция расчета радиолинии (тип 2,3,4).

Рассмотрим функционирование алгоритма.

На начальном этапе функционирования алгоритма пользователем-оператором ставится задача по организации радиолинии. Параллельно с этим на узлах сети, в автоматическом режиме выполняется зондирование ионосферы, обновляется информация о доступности узлов сети, уточняется максимальное количество узлов сети. После этого задаются пороговые значения отношения сигнал

шум  $q_{тр}$  и коэффициента разведзащищенности  $K_{рз\ тр}$ , обеспечивающие требуемое качество системы связи.



**Рис. 5. Функция расчета радиолинии**

Следующим шагом алгоритма уточняется информация о наличии средств радиоразведки. При возрастании вероятности работы средств радиоразведки, проводится расчет  $K_{рз}$  радионаправления.

При снижении значения коэффициента  $K_{рз}$  ниже требуемого значения  $K_{рз\ тр}$ , в цикле алгоритма, в зависимости от структурного ресурса сети, моделируются возможные варианты маршрута прохождения радиосигнала:

- через составную радиолинию с последовательной ретрансляцией сигналов;
- через радиолинию с параллельной ретрансляцией сигналов;
- через радиолинию с параллельной ретрансляцией сигналов и попеременным излучением;
- через комбинированную радиолинию с применением всех вышеперечисленных способов

В ходе моделирования радиолинии, для каждого составного участка определяется минимально допустимая мощность излучения передатчика. Участки составной радиолинии, для которых значение  $K_{рз}$  меньше порогового значения  $K_{рз\ тр}$  отсеиваются. Из оставшихся значений формируется таблица результатов.

Аналогичным образом формируются таблицы результатов расчетов для радиолиний (участка радиолинии) с использованием параллельной ретрансляции или попеременного излучения.

После анализа результатов расчетов по всем возможным вариантам, оператору предлагаются варианты составной радиолинии.

Разработанный алгоритм может быть применен для реализации автоматизированного элементами и режимами работы сети декаметровый радиосвязи с адаптивным управлением структурно-параметрическим ресурсом.

### Выводы

Предложенная модель автоматизированной сети радиосвязи с адаптивным управлением структурно-параметрическим ресурсом обеспечивает реализацию адаптивного управления структурным ресурсом с выбором оптимального количества ретрансляторов.

После оптимизации структуры радиосети обеспечивается выбор режимов работы ее элементов, а также адаптивное управление параметрами элементов радиосети в зависимости от сложившейся сигнально-помеховой обстановки.

Построение предложенной автоматизированной сети возможно на базе перспективных средств ДКМ радиосвязи после интеграции в них программно-аппаратного комплекса с

поддержкой протоколов физического, канального и сетевого уровня. Практическая реализация предложенной сети позволит повысить качество ДКМ радиосвязи.

Разработанный алгоритм может быть применен для реализации автоматизированного элементами и режимами работы сети декаметрового радиосвязи с адаптивным управлением структурно-параметрическим ресурсом.

#### **Библиографический список**

1 **Ватрухин, Е.М.** Новые возможности применения коротковолновой радиосвязи при решении боевой авиацией задач воздушно-космической обороны / Е.М. Ватрухин // Вестник Концерна ВКО «Алмаз – Антей» 2017. №2 (21). Стр. 16-19.

2 **Лузан, Ю.С.** Адаптивная радиосвязь в ДКМ диапазоне частот. Современное состояние и перспективы развития / Ю.С. Лузан, Н.П. Хмырова // Техника радиосвязи, 2008. Вып. 13. Стр. 3-24.

3 **Шадрин, Б.Г.** Повышение помехоустойчивости передачи данных в ведомственных сетях связи декаметрового диапазона / Б.Г. Шадрин, Д.Е. Зачатейский, В.А. Дворянчиков // Техника радиосвязи, 2018. Вып. 1 (36). Стр. 7-19.

4 MIL-STD-188-141B. Department of Defense Interface Standard, 1999.

5 **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.

6 **Рябов, А.В.** Сеть ДКМ радиосвязи с адаптивным управлением структурно-параметрическим ресурсом / А.В. Рябов, А.М. Карпель // Сборник материалов XXVI Международной научно-технической конференции «Информационные системы и технологии ИСТ-2020». Н.Новгород: НГТУ, 2020. С. 278-282.

7 **Олифер, В.Г.** Компьютерные сети. Принципы, технологии, протоколы: учебник для вузов. 3-е изд. / В.Г. Олифер, Н.А. Олифер. – СПб.: Питер, 2006. 958 с.

8 **Рябов, А.В.** Алгоритм расчета показателей разведывательной защищенности сетей авиационной радиосвязи / А.В. Рябов, М. Ю. Пресняков // Информатика: проблемы, методология, технологии: сб. материалов XIX международной научно-методической конференции / под ред. Д. Н. Борисова; Воронеж, ВГУ, 14-15 февраля 2019 г. Воронеж : Издательство «Научно-исследовательские публикации» (ООО «Вэлборн»), 2019. –2169 с. С. 521-525

## К ВОПРОСУ УЧЕТА ВЗАИМОСВЯЗИ ПРОСТРАНСТВЕННЫХ И ВРЕМЕННЫХ ФЛУКТУАЦИЙ УРОВНЯ СИГНАЛА ПРИ ОЦЕНКЕ НАДЕЖНОСТИ ПОДВИЖНОЙ НАЗЕМНОЙ РАДИОСВЯЗИ

Ю.М. Туляков

ORCID: 0000-0003-0689-0078 e-mail: [yu.m.tulyakov@rambler.ru](mailto:yu.m.tulyakov@rambler.ru)

Нижегородский государственный технический университет им. Р.Е. Алексеева,  
Волго-Вятский филиал Московского технического университета связи и информатики

**Цель работы:** За счет существующей и дополнительно исследованной взаимосвязи пространственных и временных флукуаций уровневой микроструктуры радиосигналов (со сложным многолучевым распространением) упростить и конкретизировать методику оценки надежности подвижной наземной связи.

**Методология:** Путем сопоставления пространственных и временных флукуаций уровней радиосигналов со сложным многолучевым распространением, характерных для участков, соответствующих уровневой микроструктуры, и определения стационарности этих флукуаций и возможности их оценки только по пространству (без учета временных флукуаций).

**Результаты:** Полученные обоснования оценки флукуаций радиосигналов позволили их применить для упрощения и конкретизации методики определения надежности подвижной наземной радиосвязи.

**Последствия исследования:** Конкретизирована и упрощена методика оценки надежности подвижной наземной радиосвязи.

**Ценность решения:** За счет ограниченного (пространственного) подхода к оценке флукуаций уровня сигналов упрощена и конкретизирована методика определения надежности подвижной наземной связи.

**Ключевые слова:** флукуации, уровневая микроструктура, радиосигналы, надежность связи.

## ON THE QUESTION OF ACCOUNTING FOR THE RELATIONSHIP OF SPATIAL AND TIME FLUCTUATIONS OF THE SIGNAL LEVEL WHEN ASSESSING THE RELIABILITY OF MOBILE TERRESTRIAL RADIO COMMUNICATION

Yu.M. Tulyakov

Nizhny Novgorod State Technical University. R.E. Alekseev. Nizhny Novgorod  
Volga-Vyatka Branch of Moscow Technical University of Communications and Informatics  
Nizhny Novgorod, Russia

**The purpose of the work:** Due to the existing and additionally studied relationship of spatial and temporal fluctuations of the level microstructure of radio signals (with complex multipath propagation), simplify and specify the methodology for assessing the reliability of mobile terrestrial communications.

**Methodology:** By comparing the spatial and temporal fluctuations in the levels of radio signals with complex multipath propagation, characteristic of areas corresponding to the level.

**Results:** The justifications obtained for assessing the fluctuations of radio signals made it possible to apply them to simplify and specify the methodology for determining the reliability of mobile terrestrial radio communications.

**Consequences of the study:** The methodology for assessing the reliability of mobile terrestrial radio communications has been specified and simplified.

**Value of the solution:** Due to the limited (spatial) approach to assessing signal level fluctuations, the methodology for determining the reliability of mobile terrestrial communications has been simplified and specified.

**Key words:** fluctuations, level microstructure, radio signals, communication reliability.

Надежность связи – это способность системы (канала) связи сохранять требуемое качество передачи информации, определяемое заданной помехоустойчивостью передаваемых сигналов - созданием необходимого отношения сигнал/помеха в течении заданного интервала времени, заданных условиях и территории действия системы. (При этом подразумевается, что аппаратура работает исправно). В отличие от систем с проводными линиями связи оценка надежности связи в системах, использующих в качестве линий связи **радиосигналы** – электромагнитные волны и именуемых как **радиосистемы**, значительно усложняется. Усложнение этой оценки, правильнее называя ее как, надежность **радиосвязи**, объясняется тем, что, как правило, уровни радиосигналов и помех могут меняться во времени и по зоне действия радиосистемы. Главным образом это связано с особенностями характеристик распространения радиоволн, вызывающими флуктуации уровня радиосигналов [1,2]. Учитывая популярность и перспективность систем (сетей) подвижной наземной, и в том числе сотовой радиосвязи, дадим оценку характеристикам распространения и изменения уровня радиосигналов для таких видов радиосвязи.

В радиосвязи могут использоваться электромагнитные (радио) волны (ЭМВ) различных диапазонов частот. В настоящее время наибольшее применение получили (в основном из-за бурного развития подвижной наземной связи) ЭМВ диапазонов ОВЧ (очень высоких частот -- 30 - 300 МГц) и УВЧ (ультравысоких частот – 300-3000 МГц). Также наметилась перспектива использования диапазона СВЧ (сверхвысоких частот: 3 - 30 ГГц) и более высоких частот. Как известно, ЭМВ этих частот в наземных условиях их распространения (с возможными соизмеримостями и несоизмеримостями размеров препятствиями распространения с длинами этих волн) на основании принципов Гюйгенса Френеля могут рассматриваться как ЭМВ с распространением лучем.

Расчет характеристик (трассы) распространения таких ЭМВ (радиосигналов) в свободном пространстве и в «равнинных условиях» достаточно хорошо разработан (например, в виде решений соответствующих тригонометрических задач) [2]. При передаче таких радиосигналов в условиях наличия препятствий (разного рода застроек и лесных массивов, в пригородных зонах, городских условиях с интенсивной застройкой и т.п.) распространение волн приобретает многолучевой характер с переотражением от этих препятствий и каких-либо движущихся объектов) [2]. Для такого многолучевого характера расчет характеристик распространения радиоволн усложняется и состоит в следующем:

- рассчитывается средний уровень радиосигнала в зависимости от расстояния -  $R$  до радиопередающего устройства может производиться на основании ряда с использованием эмпирических функций (моделей) для прогнозирования среднего уровня поля;

- оцениваются (определяются) возможные случайные отклонения от этого среднего значения (математического ожидания), которые могут происходить во времени и по пространству (т.е. в ближайших точках перемещения около точки с расстоянием  $R$ ).

Как показывают экспериментальные исследования [3], такие случайные отклонения от среднего значения хорошо согласуются с их логарифмически нормальным распределением, или с нормальным распределением величины отклонений, измеряемых в дБ. При существующих двух методах оценки уровня распространяемых радиосигналов: величиной электрической составляющей электромагнитного поля (ЭМП)  $-E$  или уровнем мощности ЭМП  $-P$ , такие случайные отклонения уровня ЭМП от средних значений представляются в виде [4]:

- по напряженности -  $\Delta E' = 20 \lg (\Delta E / E_M)$ , дБ<sub>E</sub>;

- по мощности-  $\Delta P' = 10 \lg (\Delta P / P_M)$ , дБ<sub>P</sub>,

где  $E_M$  и  $P_M$ , - средние значения (математические ожидания) соответственно напряженности и мощности ЭМП;  $\Delta E$  и  $\Delta P$  - абсолютные величины отклонения от средних значений уровня соответственно напряженности и мощности ЭМП.

В свою очередь сами средние значения напряженности и мощности ЭМП могут представляться в логарифмическом виде соответственно  $E_M' = 20 \lg(E_M/E_0)$ , дБ<sub>Е</sub> и  $P' = 10 \lg(P_M/P_0)$ , дБ<sub>Р</sub>, где величины  $E_0$  и  $P_0$ , относительно которых сравниваются значения напряженности и мощности ЭМП. Например, при  $E_0 = 1 \text{ мкВ/м}$  будем иметь  $E_M' = 20 \lg(E_M/E_0)$  дБ<sub>Е</sub> =  $20 \lg(E_M/1)$  дБ<sub>мкВ/м</sub>, а при  $P_0 = 1 \text{ мВт}$  имеем  $P_M' = 10 \lg(P_M/P_0)$ , дБ<sub>Р</sub> =  $10 \lg(P_M/P_0)$ , дБ<sub>мВт</sub>.

На основании выше указанных особенностей оценки величины уровней радиосигналов прогнозирование его величины в точке приема на заданном расстоянии до радиопередатчика (базовой станции-БС) можно представить в следующем виде

$$E' = E'_M \pm \Delta E'(l, p_{\phi l}) \pm \Delta E'(t, p_{\phi l}), \text{дБ}_E, \quad (1)$$

где:  $E'_M$  – математическое ожидание (среднее значение  $-E'_{\text{ср}} = E'_M$ ) уровня радиосигнала;  $\Delta E'(l, p_{\phi l})$  – случайные отклонения уровня сигнала от  $E'_M$  при перемещении по пространству ( $l$ ) на заданной высоте и фиксированном расстоянии  $R$  от БС, соответствующие вероятности появления этой величине отклонений –  $p_{\phi l}$ ;  $\Delta E'(t, p_{\phi l})$  – случайные отклонения уровня сигнала от  $E'_M$  происходящие по времени ( $t$ ), в фиксированной точке приема с заданными высотой и расстоянием  $R$  от БС, соответствующие вероятности приема с заданными высотой и расстоянием  $R$  от БС, соответствующие вероятности появления этой величине отклонений –  $p_{\phi l}$ .

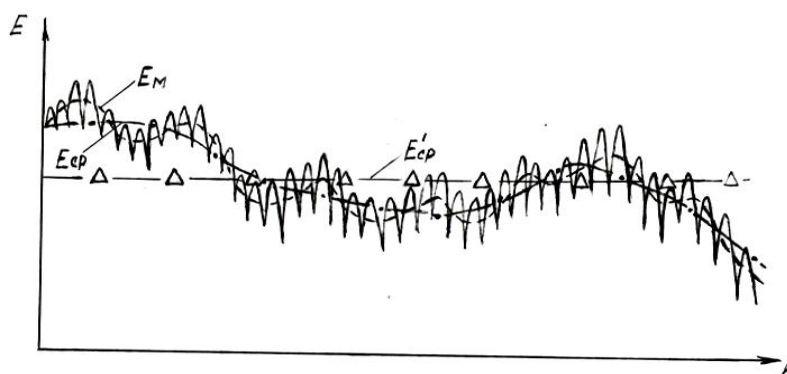
Аналогичным образом можно представить уровень сигнала по мощности [5]

$$P' = P'_M \pm \Delta P'(l, p_{\phi l}) \pm \Delta P'(t, p_{\phi l}) = \text{УММС} \pm \Delta P'(l, p_{\phi l}) \pm \Delta P'(t, p_{\phi l}), \text{дБ}_M, \quad (2)$$

где  $P'_M = \text{УММС}$  - усредненная медианная мощность сигнала (величина, которая определяется для 50% времени наблюдения и 50% точек приема, находящихся на исследуемом расстоянии  $R$ ), или в соответствии с (1) - это математическое ожидание уровня мощности радиосигнала;  $\Delta P'(l, p_{\phi l})$  и  $\Delta P'(t, p_{\phi l})$  по аналогии с (1) - это случайные отклонения уровня мощности сигнала от  $P'_M$  по времени и по пространству, соответствующие  $p_{\phi l}$  - вероятности появления этим величинам отклонений.

Как отмечалось выше, случайные величины  $\Delta E$  или  $\Delta P$  могут определяться по функциям распределения их значений для принимаемых величин  $p_{\phi l}$ . А значения математических ожиданий в месте приема радиосигналов на заданном расстоянии могут определяться (прогнозироваться) с помощью эмпирических моделей (функций), разработанных для различных условий распространения и диапазонов радиоволн. Можно перечислить ряд таких моделей для радиоволн диапазонов ОВЧ, УВЧ и СВЧ при сложной многолучевой структуре распространения: модели Бардина – Дымовича [6] или модели Окамура, Окамура - Хата, Международного союза по электросвязи (МСЭ) и др. (см. например, в [5]).

Для представления о случайных изменениях (флуктуациях) уровня этих радиоволн на рис. 1 показана типовая картина изменения уровня ЭМВ при перемещениях по пространству на определенных уровне от земли и расстоянии от радиопередатчика (БС). Этот рисунок наглядно иллюстрирует изменения уровня указанных диапазонов ЭМВ при сложной многолучевой структуре распространения, характерной для городских условий.



**Рис.1. Типовая картина изменения уровня ЭМВ сложной многолучевой структурой распространения**

Характер и параметры флуктуаций уровня ЭМП определяются экспериментально-исследовательским путем для различных условий распространения радиоволн, указанных частотных диапазонов. Базируясь на рис.1, эти исследования показывают, что флуктуации наилучшим образом описываются в виде суммы трех уровневых составляющих ЭМП [3]:

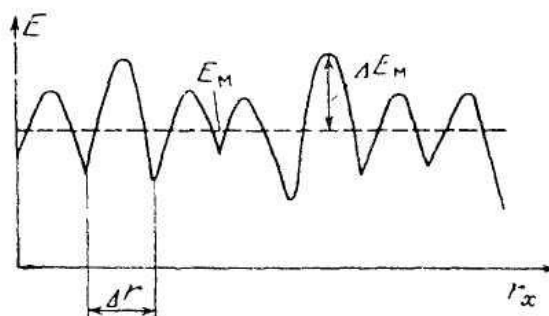
- **микроструктуры**, когда флуктуации проявляются относительно среднего (медианного) значения уровня электрической составляющей ЭМП -  $E_m$  на участках, протяженностью единиц-десятков десятков длин полуволны ЭМВ -  $(1...10) \cdot 10 \cdot (\lambda/2)$  (иногда такие флуктуации называют «быстрыми»);

- **промежуточной структурой** изменения уровня ЭМВ, которые можно рассматривать как случайное изменение медианного значения микроструктуры  $E_m$  относительно средней величины  $E_{cp}$  на участках протяженностью десятки-сотни метров, которая обусловлена интенсивностью застройки, шириной улиц и т.д.;

- **макроструктурой**, отражающей изменения среднего значения промежуточной структуры  $E_{cp}$  относительно среднего значения макроструктуры  $E_{cp, макро}$ , которые обусловлены изменением рельефа местности и группированием крупных застроек (микрорайонов) и т.д. Обычно  $E_{cp, макро}$  — определяется на участках в несколько километров, например по окружности с заданным расстоянием - радиусом от БС.

Как показывают исследования, все эти структуры флуктуаций подчиняются логарифмически-нормальным законам распределения [3].

Коррекция и уточнения для **промежуточной** структуры и **макроструктуры** могут быть сделаны с учетом реалий распространения радиоволн — учетом интенсивности застройки на заданной территории или рельефа местности и т.п. Особенно эффективность этих уточнений может быть достигнута при использовании «оцифрованных» карт исследуемой местности [7]. Для **микроструктуры** такие коррекции практически невозможны, когда флуктуации уровня поля приходится рассматривать как чисто случайный процесс.



**Рис. 2. Уровневые флуктуации микроструктуры ЭМВ**

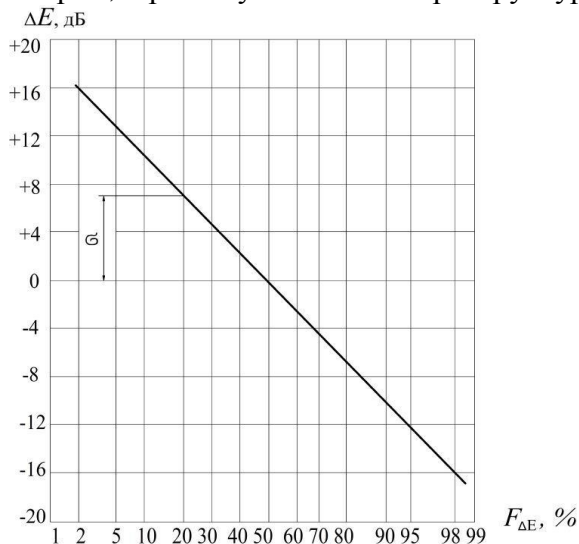
Типовой вид такой пространственной уровневой микроструктуры электрической составляющей ЭМВ  $E$  на участке  $r_x$  (по пространству) с постоянным медианным уровнем  $E_M$  и случайными отклонениями величины сигнала  $\Delta E_M$  от медианного значения, показан на рис. 2.

Заметим, что чередование минимумов величины сигнала происходит через расстояния  $\Delta r_x \approx \lambda/2$ - половине длины волны (не экстремальные изменения уровня поля имеют место для отдельных направлений в плоскости рассматриваемой микроструктуры, и периодичность повторения этих изменений может происходить через расстояния  $r \approx (0,4...0,8)\lambda/2$ ). Т.е. приближенно эти изменения можно рассматривать в виде «квазистоячих волн». Экспериментально установлено [3], что распределение величины «быстрых флуктуаций»  $20\lg(\Delta E_M/E_M)$  при этом подчиняются нормальному закону со среднеквадратическим отклонением  $\sigma_{20\lg(\Delta E_M/E_M)}$  [дБ], величина которого должна оцениваться специально для каждого конкретного случая и условий распространения радиоволн.

В итоге, с учетом рассмотренных составляющих флуктуаций уровня ЭМВ, определяемых микроструктурой, промежуточной и макро – структурами, природа которых определяется самостоятельными взаимно независимыми факторами, со своими логарифмически - нормальными законами распределения, можно представить «суммарным» законом распределением с параметрами: математическим ожиданием  $M_{\Delta E_\Sigma} = 0$  дБ и среднеквадратическим отклонением

$$\sigma_{\Delta E_\Sigma} = \sqrt{\sigma_E^2 + \sigma_{E_M}^2 + \sigma_{E_{CP}}^2}, \text{ дБ}, \quad (3)$$

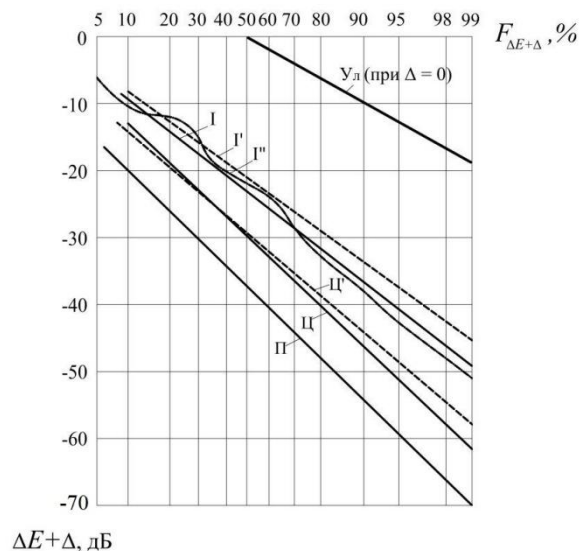
где  $\sigma_E = \sigma_{20\lg(\Delta E_M/E_M)}$ ;  $\sigma_{E_M}$ ;  $\sigma_{E_{CP}}$  – среднеквадратические отклонения, выраженные в децибелах соответственно для микро -, промежуточной и макроструктур.



**Рис.3. Распределение случайной величины  $\Delta E = \Delta E_{y,l} = \Delta E_\Sigma$  в «нормальном» масштабе**

В качестве примера, на рис. 3 показано в «нормальном» масштабе распределение случайной величины  $\Delta E = \Delta E_\Sigma$  ЭМВ диапазонов ОВЧ и УВЧ на улицах города с высокой интенсивностью застройки (при  $\sigma_{\Delta E} = \sigma_{\Delta E_\Sigma} \approx 7,8$  дБ).





**Рис.4. Функции распределения пространственных флуктуаций уровня поля ОВЧ, проникающего в помещения зданий города:  $\Pi$  – для подвальных помещений;  $\zeta$  – для цокольных этажей;  $I$  – для первых этажей; Ул (при  $\Delta=0$ ) - распределение для улиц крупных городов РФ;  $\zeta'$   $I'$   $I''$  –распределения полученные рядом исследователей зарубежем.**

На рис. 4 показаны функции распределения  $F_{\Delta E+\Delta}$  случайной величины  $\Delta E = \Delta E_{\Sigma} = \Delta E_{\text{ул}+\Delta}$  ЭМВ диапазонов ОВЧ в помещениях зданий, учитывающие случайные совместные изменения уровня радиосигнала на улице  $-\Delta E_{\text{ул}}$  и затухания уровня при проникновении в различные помещения зданий  $-\Delta$ .

Исследования пространственных флуктуаций в помещениях зданий показали [8], что в основном они характеризовались только микроструктурой (этого также следовало ожидать из-за указанных выше требований к размерам территории, на которых определяются микроструктуры). Дополнительно можно указать, что иногда, как это наглядно наблюдается из рис.1, и как указывалось выше, такие микроструктурные изменения называют «быстрыми» (хотя, как будет следовать из последующих оценок, такой термин не совсем корректен).

Таким образом итоговой оценкой уровня радиосигнала в месте приема являются медианная величина  $E_M$  и случайные флуктуации  $\Delta E = \Delta E_{\text{ул}}$  или  $\Delta E = \Delta E_{\Sigma} = \Delta E_{\text{ул}+\Delta}$ , характеризующиеся уровнем *микроструктурой*, которая в основном и определяет уровень и его возможные изменения в условиях приема радиосигналов. При этом медианная (средняя) величина  $E_M$  может прогнозироваться в соответствии с выбранной моделью (по эмпирическим формулам) [5], а случайные отклонения от этой величины  $-\Delta E$  определяются на основании экспериментальных исследований, например, для отдельных случаев, приводимых в [3,5].

Теперь остановимся на *флуктуациях уровня радиосигналов во времени*  $-\Delta E'(t, p_{\phi, \Gamma})$  [см.(1)]. Эти изменения можно наблюдать, например, в фиксированной точке или при перемещениях в месте приема сигналов. Объяснение этому можно дать с позиции рассмотренных пространственных флуктуаций, и конкретно на основе уровневой микроструктуры. Не останавливаясь на подробностях таких характеристик как стационарность и эргодичность пространственно-временных взаимосвязей флуктуаций, дадим анализ результатам ряда экспериментальных исследований в диапазонах ОВЧ и УВЧ[3]. Эти исследования показывают, что именно пространственная микроструктура определяет временные флуктуации, которые проявляют себя следующим образом:

- в фиксированной точке - как качание пространственной микроструктуры на заданном уровне точки приема (это объясняется возникающими отражениями радиоволн от движущихся объектов около точки приема – транспорта, пешеходов и т.д.);

- при перемещении, пешком или на каком - либо транспорте временные флуктуации проявляются как изменения уровня сигнала, практически полностью повторяющие пространственную уровневую микроструктуру (безусловно, это наблюдается на участках, длина которых соответствует длине, где проявляется чисто микроструктура (без влияния промежуточной структуры)).

Касаясь вопроса длительности наблюдаемых временных флуктуаций, можно привести следующие результаты экспериментальных данных. Исследования в фиксированной точке показали, что для диапазонов ОВЧ и УВЧ эти флуктуации как на улицах, так и в помещениях зданий города имеют довольно медленный характер с временем корреляции  $\tau_{\phi} \geq 0,1 \dots 1$  с [3]. Время (или частота) флуктуаций уровня радиосигнала в перемещаемой точке в большей степени будет определяться пространственной уровневой микроструктурой, и зависеть от скорости перемещения  $-V$ .

Полезно сопоставить длительность изменения (флуктуаций) уровня ЭМП с значениями длительности элементарных радиосигналов -  $\tau_{cl}$  в современных цифровых системах связи. Под  $\tau_{cl}$  будем понимать длительность элементарного радиосигнала, который может быть сформирован различными видами модуляции (и в том числе «многопозиционными»), когда скорость манипуляции (передачи) может определяться  $V \approx 1/\tau_{cl}$ .

Для современных цифровых радиосистем связи условно можно принять  $0,1 < V < 100$  Мбит/с и подсчитать  $0,001 < \tau_{cl} < 10$  мкс. Сопоставляя эти величины со значениями  $\tau_{\phi}$ , можно с уверенностью утверждать, что возможные изменения уровня радиосигнала в фиксированной точке могут проявить себя лишь после прохождения в наихудшем случае более чем через  $10^7$  элементарных сигналов (т.е. флуктуации имеют **медленный** характер).

При передвижении частота изменений уровня радиосигнала, согласно пространственно уровневой микроструктуры (рис.1) будет происходить через расстояния  $\Delta r_x \approx \lambda/2$ . В зависимости от скорости перемещения  $-V$  частота изменений

$$f_{\phi} = V/(\lambda/2). \quad (4)$$

В качестве примера, примем радиочастоту УВЧ диапазона  $f_{pч}$  -1000 МГц с длиной волны  $\lambda=0,3$  м и дадим оценку  $f_{\phi}$  для интервала скоростей от 5км/ч ( $\approx 1,4$  м/с) до 100км/ч ( $\approx 28$ м/с). В результате расчетов получаем  $9,33 < f_{\phi} < 186$  Гц. Сопоставляя эти величины с  $\tau_{cl}$  для современных цифровых радиосистем можно также прийти к выводу о **медленном** характере флуктуаций уровня радиосигнала.

Такие выводы о **медленном характере** флуктуаций (медленных «мультипликативных помехах») и определяющем влиянии пространственных флуктуаций на случайные изменения уровня радиосигналов позволяют конкретизировать оценку надежности радиосвязи. Напомним определение надежности радиосвязи:

- исходным условием является требование к допустимой вероятности ошибки передачи – приема элементарного радиосигнала-  $p_1 \leq p_{1доп}$ ,

-которое в свою очередь, в зависимости от вида радиосигнала (модуляции) определяет отношение сигнал/помеха =  $\rho$ , и с учетом помеховых условий, в итоге - необходимый минимально допустимый (пороговый) уровень сигнала  $E_{co}$  в зоне действия системы (например, для сотовой системы GSM (2G) [4] для пограничной зоны требуемый уровень по мощности должен составлять  $P' = 90$  дБм).

Зная  $E_{co}$ , надежность связи может быть оценена вероятностью  $p$  превышения (или равенства) этого порогового уровня уровнем сигнала в канале. При известных плотностях распределения вероятностей  $W(E)$  уровня сигнала надежность будет определяться (например, в процентах) интегральной функцией

$$S = p(E \geq E_{co}) \cdot 100 = 100 \int_{E_{co}}^{\infty} W(E) dE, \% . \quad (5)$$

В (5) величина  $p$  именно и определяется вероятностью (см. (1)) появления возможной величиной отклонения уровня радиосигнала от средней (медианной) его величины в зоне приема (действия) радиосистемы –  $p_{\text{ФЛ}}$ , т.е.  $p = p_{\text{ФЛ}}$ . И надежность радиосвязи  $S$  будет определяться распределениями величин  $\Delta E_{\text{ул}}$  или  $\Delta E_{\text{ул}+\Delta}$  по интегральным функциям  $F_{\Delta E}$  или  $F_{\Delta E+\Delta}$  (для сигналов ОВЧ и УВЧ см. рис. 4). Таким образом, надежность связи можно рассматривать как  $S = F_{\Delta E}$  или  $S = F_{\Delta E+\Delta}$ .

Важным результатом настоящей работы является также обоснование того, что для оценки флуктуаций уровня радиосигнала достаточно ограничиться учетом только пространственных флуктуаций, пренебрегая временными флуктуациями, и выражение (1) может быть представлено в виде:  $E' = E'_M \pm \Delta E'(l, p_{\text{ФЛ}}) \pm \Delta E'(t, p_{\text{ФЛ}}) \approx E'_M \pm \Delta E'(l, p_{\text{ФЛ}})$  дБ<sub>Е</sub>.

Этим результатом можно разъяснить данные, приводимые по флуктуациям уровня радиосигналов СВЧ диапазона в [5], где указывается некая средняя величина дисперсии флуктуаций, не оговаривая наличие временных флуктуаций.

*Замечание:* Выше представлены результаты исследования характеристик распространения радиоволн со сложной многолучевой структурой, позволяющие дать оценку изменениям (флуктуациям) уровня радиосигналов в простейшем (классическом) варианте – путем суммирования лучей распространения в месте приема. Другим внедряемым сейчас способом (технологией) приема радиосигналов с осуществлением переориентации произвольно направленных лучей радиоволн в лучи, которые синфазно складываются с целью повышения уровня принимаемого сигнала. Одним из примеров такой обработки принимаемых радиосигналов является технология *Rake*, применяемая в CDMA и WCDMA (Wideband Code Division Multiple Access- (3G)) сотовых системах. Несмотря на отличие вышерассмотренных характеристик уровней радиоволн для произвольно суммируемых лучей от суммирования по технологии *Rake*, результаты, представленные в данной работе, могут быть полезными для представления об энергетике радиосигналов в месте приема и определения сравнительной эффективности использования *Rake* радиоприема.

### Выводы

1. Надежность связи в радиосистеме подвижной связи в основном определяется пространственно - временными флуктуациями уровня радиосигналов в зоне действия радиосистемы.
2. При сложной многолучевой структуре распространения радиоволн, свойственной условиям города, эти флуктуации подчиняются логарифмически - нормальным законам распределения, характеризуя случайные изменения уровня радиосигнала для двух видов условий: на улицах и в помещениях зданий.
3. По величине флуктуаций их характеризуют тремя видами пространственных уровневых структур, наиболее важной из которых при определении надежности связи является микроструктура, определяющая медленные флуктуации уровня радиосигнала.
4. Поскольку временные флуктуации уровня радиосигналов определяются пространственной уровневой микроструктурой, то для определения надежности радиосвязи достаточно ограничиться учетом только пространственной уровневой микроструктуры, не учитывая временные флуктуации.

### Библиографический список

1. **Долуханов, М.П.** Распространение радиоволн. – М.:Связь, 1972. – 336 с.
2. **Туляков, Ю.М.** Пространственная надежность прохождения радиосигналов со сложной многолучевой структурой распространения в условиях города (на улицах и при проникновении в помещения). // Вестник нижегородского университета им. Н.И. Лобачевского, №5, часть 1, 2010. - С.75-84.
3. **Туляков, Ю.М.** Системы персонального радиовызова.– М.: Радио и связь, 1988–168с.
4. **Туляков, Ю.М.** О децибелах. М. Горячая линия-Телеком, 2019 г.-106 с
5. **Маковеева, М.М., Шинаков, Ю.С.** Системы связи с подвижными объектами: Учебное пособие для вузов. – М.: Радио и связь, 2002. – 440 с.
6. **Бардин, Н.И., Дымович, Н.Д.** Распространение ультракоротких радиоволн в условиях крупного города // Электросвязь. 1964. -№7. – С. 15-18.
7. **Милютин, Е.Р., Василенко, Г.О.** Повышение точности расчета ослабления поля с помощью калибровки и цифровых карт местности //Электросвязь.2004.- № 2 – С. 38-40.

**РАСШИРЕНИЕ КЛАССА ЛИНЕЙНЫХ РЕЗОНАНСНЫХ СИСТЕМ ПРИ  
АППРОКСИМАЦИЯ СВОБОДНОГО ПРОЦЕССА ОРТОГОНАЛЬНЫМИ  
ПОЛИНОМАМИ**

**С.С.Зельманов**

**ORCID: 0000-0002-6634-942X e-mail: zelmanss@yandex.ru**

Нижегородский государственный технический университет им. Р.Е.Алексеева

**Цель работы:** Она состоит в доказательстве возможности расширения класса резонансных линейных систем за счет систем, корни характеристических уравнения которых являются отрицательными действительными числами.

**Методология:** С использованием выражение свободного процесса резонансной линейной динамической системы решается задача аппроксимации этого свободного процесса резонансной системы в ортогональном базисе с использованием процедуры Грама-Шмидта.

**Результаты:** Полученное выражение определяет структуру резонансной системы с экспоненциальными собственными процессами. Это отличает их от систем с гармоническими собственными процессами. Эти системы используют только один накопитель энергии.

**Последствия исследования:** При анализе линейных систем второго порядка исключат аperiодические и критические решения, поскольку все эти решения будут колебательными разной интенсивности.

**Ценность решения:** Наличие экстремума в спектре подтверждает резонансный характер процесса в системе, корни характеристического уравнения которой являются отрицательными действительными числами, а некомплексно-сопряженными. Это и означает расширение класса резонансных систем.

**Ключевые слова:** ортогонализация, процедура Грама-Шмидта, отрицательные корни,

**EXPANSION OF THE CLASS OF LINE RESONANCE SYSTEMS AT  
APPROXIMATION OF A FREE PROCESS BY ORTHOGONAL POLYNOMS**

**S.S. Zelmanov**

Nizhny Novgorod state technical university. R.E. Alekseev

**Purpose:** It consists in proving the possibility of extending the class of resonant linear systems by adding to this class systems, whose characteristic equation roots are negative real numbers.

**Methodology:** Using the free process expression of a resonant linear dynamical system, the problem of approximating this free process of a resonant system in an orthogonal basis is solved using the Gram-Schmidt procedure.

**Results:** The resulting expression determines the structure of a resonant system with exponential eigenprocesses. But a harmonic free and weakly expressed process. This distinguishes them from traditional systems with harmonic eigenprocesses. Structurally, such systems are implemented using only one energy storage device.

**Consequences of the study:** In the analysis of linear systems of the second order, aperiodic and critical solutions will be excluded, since all these solutions will be oscillatory different intensity.

**Value of the solution:** The presence of an extremum in the spectrum confirms the resonant nature of the process in the system, the roots of the characteristic equation of which are negative real numbers, in contrast to the traditional system, which has complex conjugate roots. This means an extension of the class of resonant systems.

**Keywords:** orthogonalization, Gram-Schmidt procedure, negative roots.

## 1. Цель работы и постановка задачи

Цель состоит в доказательстве возможности расширения класса резонансных линейных систем за счет добавления к ним систем, корни характеристических уравнений которых являются отрицательными действительными числами. При этом анализ линейных систем второго порядка должен исключить традиционные аperiodические и критические решения, поскольку все эти решения, строго говоря, являются колебательными, но малой интенсивности, как недостаток, легко устранимой. Традиционные представления о признаках резонанса были заложены достаточно давно и отвечали требованиям своего времени [1]. Физические соображения о том, что в системах второго порядка резонанс при значительных потерях может быть почти «не заметен» позволили существовать в литературе аperiodическим решениям, не затрагивающим поведение тока в таких системах. Однако, «не заметен» это вовсе не значит, что «не существует». Как выяснилось, учёт влияния потерь в определённых резонансных системах привел к заключению, что потери играют в них большую конструктивную роль. [2].

Структура резонансных линейных систем определяется характером их свободного процесса, т.е. импульсной характеристикой, которая может быть аппроксимирована различными способами, одним из которых является способ точечной или узловой аппроксимации свободного процесса системы.

Недостатком этого способа является то, что для его реализации требуется решение достаточно большого числа уравнений. Другим недостатком является некоторая неопределенность в поведении аппроксимирующей функции в интервалах между узлами аппроксимации.

Поэтому задача состоит в проведении более точной аппроксимации, которая может быть обеспечена за счет использования ортонормальной системы экспоненциальных функций. Ортогональность этих функций, как высшая степень их линейной независимости, обеспечивает требуемую точность аппроксимации.

При этом результат аппроксимации, открывает возможность формирования систем с различными свойствами, в том числе и резонансными.

## 2. Ортогонализация линейно независимых экспоненциальных функций

Для получения ортогонального базиса экспоненциальных функций используется процедура Грама - Шмидта [3]. Следует заметить, что на эту возможность указал Л.Фрэнкс в работе [4], замечая, что на каждом  $i$ -ом шаге должна увеличиваться размерность пространства, натянутого на ортогональный базис.

Рассмотрим систему линейно независимых экспоненциальных функций (1):

$$f_0(t) = e^{-\alpha_0 t} : f_1(t) = e^{-\alpha_1 t} : f_2(t) = e^{-\alpha_2 t} \dots f_n(t) = e^{-\alpha_n t} : \quad (1)$$

Алгоритм процедуры Грама-Шмидта позволяет получить ортонормальную систему функций:

$$Z_1(t); Z_2(t); \dots; Z_n(t), \quad \text{где} \quad Z_1(t) = \frac{f_0(t)}{\sqrt{\int_{-\infty}^{\infty} f_0^2(t) dt}} ; \text{ введём вспомогательную}$$

$$\text{величину } \theta_1(t) = f_1(t) - f_{11} Z_1(t), \text{ где } f_{11} = \int_{-\infty}^{\infty} f_1 \frac{f_0(t)}{\sqrt{\int_{-\infty}^{\infty} f_0^2(t) dt}} dt \text{ при } n = 1$$

$$\text{Тогда } Z_n(t) = \frac{\theta_n(t)}{\sqrt{\int_0^\infty \theta_n^2(t) dt}}; \theta_n(t) = f_n(t) - \sum_{j=1}^{n-1} f_{nj}(t) Z_j(t); \quad (2)$$

$$\theta_n(t) = f_n(t) - \sum_{j=1}^{n-1} f_{nj}(t) Z_j(t) \quad \theta_1(t) = f_1 - S_{n2} Z_2(t) - \dots - S_{n(n-1)} Z_{n-1}(t)$$

$$f_{n(n-1)} = \int_{-\infty}^{\infty} f_n(t) Z_{n-1}(t) dt \quad \text{где при } n \geq 2$$

Рассмотрим первый член  $Z_1(t)$  ортогональной системы функций.

$$\text{Если } f_0(t) \neq 0, \text{ то } Z_1(t) = \frac{f_0(t)}{\sqrt{E_0}}; \quad E_0 = \int_{-\infty}^{\infty} f_0^2(t) dt = \frac{1}{2\alpha_1}$$

$$\text{Тогда } Z_1(t) = \frac{f_0(t)}{\sqrt{E_0}} = \sqrt{2\alpha_1} e^{-\alpha_1 t} \quad (3)$$

Легко убедиться в том, что  $Z_1(t)$  выражения (3) будет иметь единичную энергию.

$$\int_{-\infty}^{\infty} Z_1^2(t) dt = \int_0^{\infty} (\sqrt{2\alpha_1})^2 e^{-2\alpha_1 t} dt = 1$$

Поскольку  $f_0(t) = Z_1(t) \sqrt{E_0}$  то «проекция» функции  $f_0(t)$  на «ось» в новом ортогональном базисе будет  $f_{10} = \sqrt{E_0}$ .

Введем вспомогательную функцию  $\theta_2(t) = f_2(t) - f_{21} Z_1(t)$  и в соответствии с алгоритмом (2) определим  $f_{21}$ :  $f_{21}(t) = \int f_2(t) Z_1(t) dt = \frac{\sqrt{2\alpha_1}}{\alpha_1 + \alpha_2}$

$$\theta_2(t) = e^{-\alpha_1 t} - \frac{\sqrt{2\alpha_1}}{\alpha_1 + \alpha_2} \sqrt{2\alpha_1} e^{-\alpha_1 t} = e^{-\alpha_2 t} - \sqrt{2\alpha_1} e^{-\alpha_1 t} = e^{-\alpha_2 t} - \frac{2\alpha_1}{\alpha_1 + \alpha_2} e^{-\alpha_1 t}$$

Тогда второй член ортонормальной системы в соответствии с (2) будет иметь вид (4):

$$Z_2(t) = \frac{\theta_2(t)}{\sqrt{\int_0^\infty \theta_2^2(t) dt}} = -\frac{2\alpha_1 \sqrt{2\alpha_2}}{|\alpha_1 - \alpha_2|} e^{-\alpha_1 t} + \frac{\sqrt{2\alpha_2} (\alpha_1 + \alpha_2)}{|\alpha_1 - \alpha_2|} e^{-\alpha_2 t} \quad (4)$$

Этот процесс может быть продолжен. В результате процесс ортогонализации функций экспоненциального ряда приводит к образованию базиса в виде ортогональных полиномов вида (5):

$$Z_1(t) = B_{11} e^{-\alpha_1 t}$$

$$Z_2(t) = B_{12} e^{-\alpha_1 t} + B_{22} e^{-\alpha_2 t}$$

$$Z_3(t) = B_{13} e^{-\alpha_1 t} + B_{23} e^{-\alpha_2 t} + B_{33} e^{-\alpha_3 t}$$

.....

$$Z_n(t) = B_{1n}e^{-\alpha_1 t} + B_{2n}e^{-\alpha_2 t} + \dots + B_{nn}e^{-\alpha_n t}$$

В дальнейшем мы будем называть их экспоненциальными ортогональными  $Z(t)$ -полиномами.

### 3. Аппроксимация свободного процесса резонансной системы в ортогональном базисе

Для решения задачи аппроксимации непрерывной функции свободного процесса резонансной системы  $u_{св.}(t)$  в ортогональном базисе необходимо определить «проекции» этого процесса на «оси» полученного ортогонального базиса (5). Иными словами, необходимо определить «амплитуды» спектральных составляющих свободного процесса в этом базисе в соответствии с выражениями (6):

$$A_1 = \int u_{св.}(t) Z_1(t) dt ; A_2 = \int u_{св.}(t) Z_2(t) dt ; \dots ; A_n = \int u_{св.}(t) Z_n(t) dt \quad (6)$$

Тогда аппроксимирующий полином свободного процесса будет иметь вид (7):

$$u_{св.}(t) = A_1 Z_1(t) + A_2 Z_2(t) + A_3 Z_3(t) + \dots + A_n Z_n(t) \quad (7)$$

После подстановки (5) в (6) и проведения преобразований получим:

$$u_{св.}(t) = C_1 e^{-\alpha_1 t} + C_2 e^{-\alpha_2 t} + \dots + C_n e^{-\alpha_n t} ,$$

где

$$C_1 = A_1 B_{11} + A_2 B_{12} + \dots + A_n B_{1n}$$

$$C_2 = A_2 B_{22} + A_3 B_{23} + \dots + A_n B_{2n}$$

$$\dots$$

$$C_n = A_n B_{nn}$$

Очевидно, что точность аппроксимации свободного процесса будет определяться количеством экспонент, которые будут использоваться в выражении (7). Это выражение будет определять также структуру самой системы.

Применим рассмотренную методику для ортогонализации следующей системы экспоненциальных функций с кратными показателями (8):

$$f_1(t) = e^{-\alpha t}; f_2(t) = e^{-2\alpha t}; f_3(t) = e^{-3\alpha t}; \dots ; f_n(t) = e^{-n\alpha t}; \quad (8)$$

Первые четыре члена ортонормального базиса будут иметь вид (9):

$$\begin{aligned} Z_1(t) &= \sqrt{2\alpha} e^{-\alpha t}; & Z_2(t) &= \sqrt{\alpha} (-4e^{-\alpha t} + 6e^{-2\alpha t}) \\ Z_3(t) &= \sqrt{6\alpha} (3e^{-\alpha t} - 12e^{-2\alpha t} + 10e^{-3\alpha t}) \\ Z_4(t) &= \sqrt{8\alpha} (-4e^{-\alpha t} + 30e^{-2\alpha t} - 60e^{-3\alpha t} + 35e^{-4\alpha t}) \end{aligned} \quad (9)$$

Графики этих  $Z$ -полиномов представлены на рис.1.

Теперь, используя  $Z(t)$ -полиномы, т.е. ортогональный базис (9), проведем аппроксимацию свободного процесса реальной резонансной системы. выражение которой в общем случае имеет вид (10):

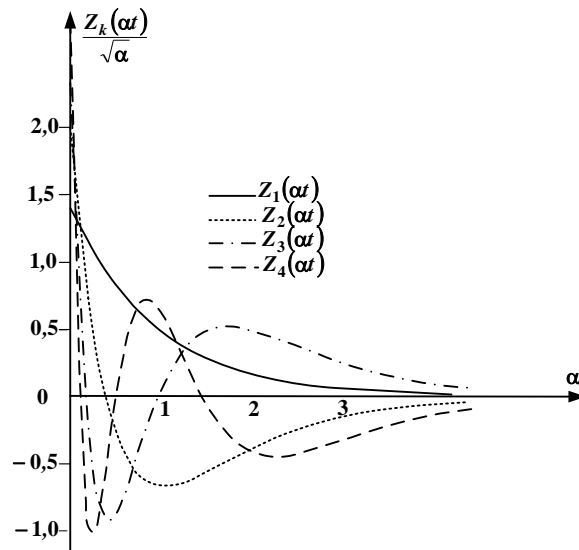
$$u_{св.}(t) = \omega_0 e^{-\alpha t} \sin \omega_0 t \quad (10)$$

Определим с использованием  $Z(t)$ -полиномов «амплитуды» двух первых составляющих «спектра» выражения (10) в ортонормальном экспоненциальном базисе при  $\omega_0 = 4\pi\alpha$ . Эти выражения будут иметь вид:



$$A_1 = \int_{-\infty}^{\infty} u_{c6}(t) Z_1(t) dt = 1,375\sqrt{\alpha}$$

$$A_2 = \int_{-\infty}^{\infty} u_{c6}(t) Z_2(t) dt = 1,776\sqrt{\alpha} \quad (11)$$



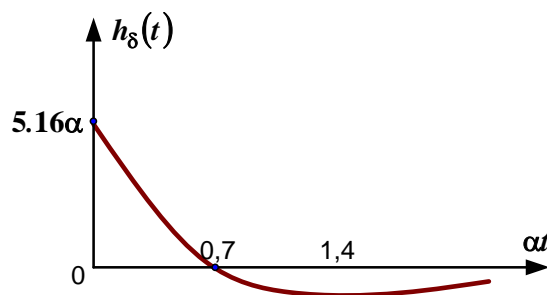
**Рис.1. Графики первых четырех Z-полиномов экспоненциального ортогонального базиса**

Тогда первое приближение аппроксимируемого свободного процесса (10) в соответствии с соотношением (7) будет иметь вид первого приближения импульсной характеристики резонансной системы:

$$h_{\delta}(t) = A_1 Z_1(t) + A_2 Z_2(t) = 1,93\alpha e^{-\alpha t} - 7,1\alpha e^{-2\alpha t} + 10,65\alpha e^{-2\alpha t} =$$

$$= 5,16\alpha(2e^{-2\alpha t} - e^{-\alpha t}) \quad (12)$$

График этого процесса представлен на Рис. 2.



**Рис. 2. График затухающего гармонического процесса (первого приближения)**

Введём обозначения:  $\alpha = \alpha_1; 2\alpha = \alpha_2$ .

Тогда полученное выражение (12) с точностью до постоянного множителя будет соответствовать решению уравнения свободного процесса тока в одноконтурной LCR-системе (13):

$$i(t) = \frac{1}{L(\alpha_2 - \alpha_1)} (\alpha_2 e^{-\alpha_2 t} - \alpha_1 e^{-\alpha_1 t}) \quad (13)$$

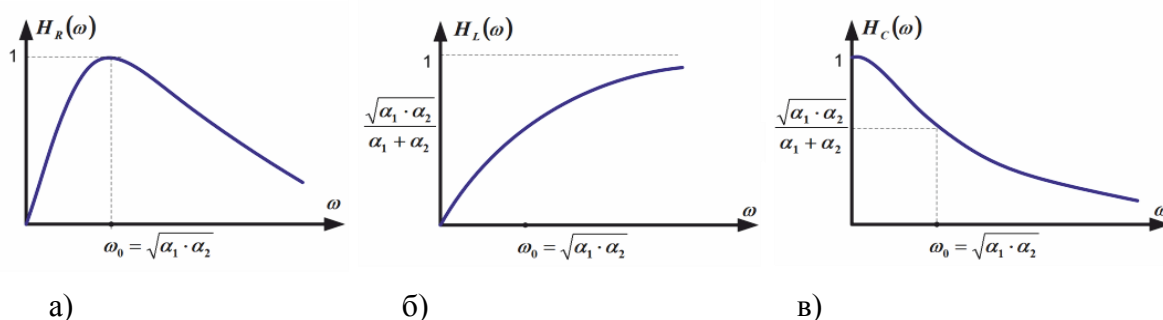
Выражение для напряжения на резисторе системы будет иметь вид выражения (14)

$$u_R(t) = iR = \frac{R}{L} \frac{1}{\alpha_2 - \alpha_1} (\alpha_2 e^{-\alpha_2 t} - \alpha_1 e^{-\alpha_1 t}) \quad (14)$$

. Выражение спектра этого свободного процесса или модуля коэффициента передачи имеет вид (15):

$$S_{\delta R}(\omega) = H_R(\omega) = \frac{2\omega\alpha}{\sqrt{(\alpha_1\alpha_2 - \omega^2)^2 + \omega^2(\alpha_1 + \alpha_2)^2}} \quad (15)$$

Выражение (13), представленное на рис.3а, имеет максимум на частоте  $\omega = \sqrt{\alpha_1\alpha_2}$



**Рис.3. Графики модулей спектров свободных процессов тока**

Наличие максимума спектра напряжения на резисторе, пропорционального току свободного процесса тока в системе, свидетельствует о резонансном поведении тока. [5].

На рис 3б и 3в для сравнения представлены спектры напряжений на индуктивности и ёмкости LCR-контра, носящие экспоненциальный характер, свидетельствующий об отсутствии резонанса. И так, в одной и той же системе резонанс есть и в то же время его как бы нет! Традиционно с явлением резонанса в линейной системе связывают наличие свободного процесса, имеющего затухающий колебательный характер. Однако, вид полученного свободного процесса, будучи представлен суммой двух экспонент (12), значительно отличается от затухающего колебания. Возникает определенное противоречие, требующее разъяснения. Если речь идет о возможности резонанса в системе, то решение ее уравнения в традиционном представлении должно иметь вид колебательного, а не аperiodического процесса. Можно ли считать напряжение, представленное на рис.2 периодическим колебанием? Подобно колебанию оно меняет знак, но однократно в момент  $\alpha t = 0,7$ , но периодичность как бы не наблюдается. Но если мы рассмотрим колебание вида  $u_1(t) = U_m e^{-\alpha t} \sin \omega_0 t$ , то, строго говоря, его тоже нельзя назвать периодическим. Однако, такой свободный процесс также соответствует резонансной системе с потерями. Периодическим в строгом смысле можно назвать только колебание вида  $u_2(t) = U_m \sin \omega t$ , соответствующее системе без потерь, т.е. идеальной системе. Рассмотренное решение (13) для тока свидетельствует о резонансе, но слабо выраженном. Однако, он может быть сделан

как угодно острым. Очевидно, что характер режима в одноконтурной системе определяется, прежде всего, поведением тока в ней.

Что касается элементов  $L$  и  $C$ , то индуктивность и емкость системы как частотно-зависимые элементы трансформируют резонансное поведение тока в соответствии со своими свойствами в экспоненциальные процессы напряжения на них, спектры которых экстремума не имеют. Поэтому характер напряжения на этих элементах может быть отличным от характера поведения тока в системе. Отсюда следует различный характер спектров свободного процесса, относящихся к этим элементам.

Система второго порядка не может, в строгом смысле, иметь экспоненциального решения, а решение как сумма экспонент может при определенных условиях представлять собой слабо выраженное колебание. Что касается корней характеристического уравнения системы, то соответственно они представляют собой отрицательные действительные числа.

#### 4. Заключение

Итак, среди резонансных линейных систем наблюдаются системы, отвечающие как гармоническим, так и негармоническим разложениям. Это означает, что свободные процессы этих систем, носящие колебательный характер различной интенсивности, могут включать в себя экспоненциальные составляющие. При этом, свободные процессы, являющиеся суммой экспонент, представляют собой колебание как результат неполной компенсации этих экспонент и в таких системах имеет место явление резонанса. Следует заметить, что потери в колебательной системе, приводящие к сумме экспоненциальных решений, для ряда систем носят конструктивный характер. К таким системам относятся, прежде всего, резонансные  $RC$ -системы, колебательный характер свободных процессов которых является суммой не двух, а большего числа экспонент [6].

Отнесение таких систем к резонансному классу линейных стационарных систем позволяет заключить о расширении класса линейных резонансных систем.

#### Библиографический список

- 1. Андронов, А.А.** Теория колебаний / А.А. Андронов, А.А. Витт, С.Э. Хайкин. – М.: ГИФМЛ, 1959
- 2. Зельманов, С.С.** Исследование явления резонанса в цепной  $RC$ -линии / С.С.Зельманов// Т•Сотт (Телекоммуникации и транспорт), Спецвыпуск по итогам 3-ей отраслевой научной конференции «Технологии информационного общества», №8, 2009.
- 3. Возенкрафт, Дж.** Теоретические основы техники связи: пер. с англ. / Дж. Возенкрафт, И. Джекобс. – М.: Мир, 1969
- 4. Фрэнкс, Л.** Теория сигналов: пер. с англ. / Л. Фрэнкс. – М.: Сов. радио, 1974.
- 5. Зельманов, С.С.** Резонанс в линейной стационарной системе с экспоненциальными собственными процессами //Известия ЮФУ. Технические науки.- Таганрог: Изд-во ТТИ ЮФУ, 2010.-12(113). №3С.116-125.
- 6. Зельманов, С.С.** Развитие теории резонанса в линейных стационарных и управляемых системах. Детектирование обобщенных АМ и ЧМ-колебаний: монография /С.С.Зельманов.//; Московский технический университет связи и информатики (Волго-Вятский филиал МТУСИ).– М., 2007.- 200 с

**ОПТИМИЗАЦИЯ ФИЗИЧЕСКОГО И КАНАЛЬНОГО УРОВНЕЙ  
МОБИЛЬНОЙ ВИДЕО СЕТИ****Т.И. Горячева****ORCID:0000-0003-4668-4156, e-mail: gor43210@yandex.ru****Я.С. Некрасов****ORCID: 0000-0003-0625-4717, e-mail: Nekrasov.yaroslav23@gmail.com****С.А. Заглумонин****ORCID: 0000-0002-5934-3134, e-mail: zaglumonin99@gmail.com**

Нижегородский государственный технический университет им. Р.Е. Алексеева

Доклад посвящен разработке адаптивной формы протокола в виде сочетания протокола множественного доступа к каналу и структуры сигнала. Он приспособляется к текущему состоянию сетевого сценария (количеству узлов сети) и к модели трафика (объёму пакета, количеству передач пакетов в секунду в одном потоке, количеству потоков от каждого узла). При изменениях сетевого сценария и/или модели трафика синтезируется новая форма волны, т.е. сначала вычисляется оптимальный протокол доступа к каналу по критерию минимума времени задержки передачи пакета, джиттера. Затем для выбранного протокола доступа синтезируется оптимальная структура OFDM мультиплекса, которая позволяет передать пакет в одном временном слоте, используя оптимальный вид модуляции M-QAM и кодирования поднесущих OFDM с учётом энергетики канала. Исследована и выбрана структура OFDM сигнала, оптимальная по критерию минимума требуемой полосы. Предложены алгоритмы расчёта адаптивных форм волны для создания баз знаний (таблиц политик), встраиваемых в мобильные узлы для ускорения выбора адаптивной формы волны по команде, содержащей номера сетевого сценария, модели трафика и момент времени перехода на новую форму. Для каждой выбранной формы волны вычисляется требуемая полоса сигнала и остаток общей полосы для гибкого использования других виртуальных сетей. Алгоритмы синтеза форм волны могут эффективно применяться в мобильных наземных и летающих видео сетях гражданского и военного назначений, реализующих менеджмент на базе политик (PBNM).

**Ключевые слова:** Адаптивная форма волны для видеосети, структура OFDM сигнала, менеджмент сетями на базе политик.

**DEVELOPMENT OF ADAPTIVE WAVEFORM FOR MOBILE VIDEO NETWORK****T.I. Goryacheva, Ya.S. Nekrasov, S.A. Zaglumonin**

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

**Purpose:** The report is devoted to the development of an adaptive waveform, which is a combination of a multiple channel access protocol and a signal structure. It adapts to the current state of the network scenario and to the traffic model.

**Design/methodology/approach:** When the network scenario and/or traffic model changes, a new waveform is synthesized, i.e. first, the optimal channel access protocol is calculated according to the criterion of minimum packet transmission delay time, jitter. Then, for the selected access protocol, the optimal structure of the OFDM multiplex is synthesized, which allows the packet to be transmitted in one time slot using the optimal type of M-QAM modulation and OFDM subcarrier coding, taking into account the channel energy.

**Conclusions:** Algorithms for calculating adaptive waveforms are proposed that can be used to create knowledge bases (policy tables) embedded in mobile nodes to speed up the selection of an adaptive waveform by a command from the manager, which should contain the network scenario number, traffic model number and the time of transition to new waveform.

**Consequences of research/limitations:** For each selected waveform, the required signal bandwidth is calculated, as well as the remainder of the total bandwidth, which can be used flexibly for other virtual networks.

**Originality/value:** Waveform synthesis algorithms can be effectively applied to civil and military mobile terrestrial and airborne video networks implementing policy-based network management.

**Keywords:** Adaptive waveform for video network, OFDM signal structure, policy-based network management (PBNM).

## Введение

Рост мобильного трафика, увеличение числа устройств, подключаемых к сети, реализация новых приложений, нехватка частотного спектра, рост требований к пропускной способности и к объёму мобильных сетей, стимулируют внедрение более высокоскоростных сигнально-кодовых конструкций (СКК), новых протоколов доступа к каналу, обеспечивающих повышение спектральной эффективности при сохранении высокой надёжности и робастности связи, новых политик управления сетями (PBNM – Policy Based Network Management) для реализации стремительных операций и разнообразных миссий.

Эти процессы находят отражение в пятом поколении мобильных сетей связи (5G), широкое внедрение которых началось в 2020 г. Высокие участки частотного спектра необходимы сетям 5G для достижения скоростей порядка 20 Гбит/с и выше [1, 4, 9]. Они позволяют реализовать 3D-видео в формате UHD (Ultra High Definition), дополненную реальность (AR), облачные сервисы, голографическую связь, тактильный интернет, самоорганизацию сетей летающих беспилотников FANET (Flying Ad hoc Network) и др.

Указанные характеристики в десятки раз превышают параметры сетей стандарта 4GLTE-Advanced (LTE-A) [3]. Для реализации требований к высокоскоростным мобильным широкополосным сетям (eMBB) внедряются новые эффективные сигнально-кодовые конструкции (F-OFDM + M-QAM) на базе фундаментального, гибкого, фильтруемого, ортогонального частотного мультиплексирования поднесущих (F-OFDM – Fundamental, Flexible, Filtered – Orthogonal Frequency Division Multiplexing) в сочетании с модуляцией каждой поднесущей M-позиционной квадратурной модуляцией (M-QAM – M-position Quadrature Amplitude Modulation). Сигнал F-OFDM пятого поколения [2] похож на сигнал четвертого поколения OFDM 4G (LTE-Advanced, 802.16m – WI-MAX), но имеет возможность гибкого назначения внутри одного сигнала (мультиплекса) разных значений ортогонального разноса по частоте между несущими, также разных значений защитных временных интервалов (циклических префиксов) для обеспечения функции передачи данных разных виртуальных сетей в одном сигнале (VNF – Virtualization Network Function) [9]. Рассмотрим подробнее, за счёт чего можно обеспечить такой прогресс в технических характеристиках и удовлетворить требования пятому поколению. Начнём с наиболее фундаментального аспекта – физического и канального уровней.

## Исследование предметной области

Чтобы выбрать оптимальную структуру мультиплекса OFDM сфокусируем внимание на структурах сигналов OFDM, рекомендуемых рабочей группой LTE (Long-Term Evolution) [3] долгосрочного развития мобильной связи, которая внедрила большое количество сотовых систем связи на базе технологии 4G-LTE и 4G-LTE-Advanced с помощью разбиения поднесущих сигнала OFDM на ресурсные блоки (РБ) по 12 поднесущих (12 полос  $\Delta F$  ортогонального частотного разноса,  $\Delta F = 1/T_U$ , где  $T_U$  – время передачи полезного сигнала и анализа его БПФ при приёме символа сигнала OFDM длительностью  $T_S = T_g + T_U$ , где  $T_g$  – защитный интервал (префикс) вводится в начале каждого символа для борьбы с задержкой многолучёвости, вызывающей глубокие частотно-селективные замирания. Ресурсный блок OFDMA системы стандарта 4G-LTE показан на рис. 1 [5].

Полоса  $W_C$  частот сигнала является переменной величиной и зависит от количества используемых поднесущих (ПН) (размерности преобразования Фурье). Разбиение на РБ реализовано для удобства масштабирования полосы сигнала путём добавления или отключения РБ. Каждый РБ обрабатывается отдельной микросхемой ПЛИС (FPGA). Ресурсные блоки могут использоваться для множественного доступа (OFDMA), когда конкретному узлу назначаются РБ с конкретными номерами внутри сигнала OFDM. В системе LTE информация передаётся кадрами длительностью 100 миллисекунд (мс). Каждый кадр состоит из 10 подкадров длительностью 10 мс, которые делятся на 20 временных слотов. Длительность слота составляет 0.5 мс, и делится на 7 OFDM символов (для

стандартного префикса) или на 6 OFDM символов (для расширенного префикса). Длина символов равна, соответственно,  $T_{S1} = 71,43$  мкс или  $T_{S2} = 83,34$  мкс.

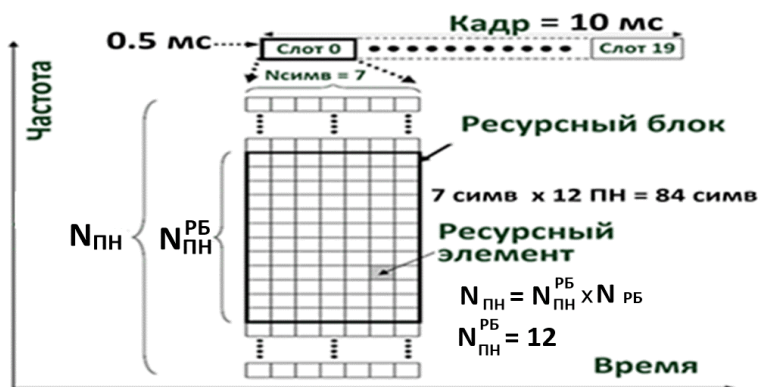


Рис. 1. Ресурсный блок OFDMA системы стандарта LTE-4G

Защитные интервалы  $T_g$  приведены в табл. 1 [5]. Их общая длительность на слот равна 36,47 мкс при 7-ми символах на слот и 100 мкс при 6-ти символах на слот. С учётом защитных интервалов  $T_g$  (циклических префиксов), вставляемых в начале каждого символа, время излучения сигнала внутри слота длиной 0,5 мс = 500 мкс равно  $T_{из} = 500 - 36,47 = 463,53$  мкс для 7 символов и 400 мкс для 6 символов на слот. Ширина полосы одной поднесущей (ПН) для стандарта 4G-LTE равна  $\Delta F = 15$  кГц. Таким образом, ширина полосы РБ ( $W_{РБ}$ ), содержащего 12 ПН, составляет  $W_{РБ} = \Delta F \cdot (N_{ПН}^{РБ}) = 15 \text{ кГц} \cdot 12 = 180 \text{ кГц}$ , а длительность  $T_{РБ} = 0,5$  мс. РБ содержит 12 ПН по 7 символов  $N_{симв}^{РБ} = 12 \cdot 7 = 84$  симв.

Таблица 1

Длительности используемых префиксов в LTE-4G

Конфигурация	Длина префикса $T_g$ , мкс	Эквивалентная разность хода лучей, км
Стандартный префикс – защитный интервал	4,69 (первый символ)	1,4
	5,21 (последующие символы) (всего 36,47 мкс на слот)	1,55
Расширенный префикс	16,67 (100 мкс на слот)	5

В табл. 2 представлены значения параметров сигнала 4G-LTE [3].

Таблица 2

Параметры сигналов OFDM системы 4G-LTE для  $\Delta F = 15$  кГц

Полоса $W_c$ частот, МГц	1,4	3	5	10	15	20
Количество доступных ресурсных блоков (РБ), $N_{РБ}$	6	15	25	50	75	100
Количество используемых поднесущих на сигнал $N_{ПН}$	72	180	300	600	900	1200

Проверим, можно ли передать таким сигналом наш видео пакет в одном слоте.

Как видно из табл. 2, в полосе 20 МГц количество поднесущих равно 1200. При использовании вида модуляции 64-QAM каждый символ переносит  $\log_2(64) = 6$  бит. Одна ПН содержит 7 символов и переносит  $7 \cdot 6 = 42$  бита. Один сигнал из 1200 ПН переносит  $42 \cdot 1200 = 50400$  бит. А нам надо передать пакет 400000 бит за время слота 0.5 мс. Следовательно, нужен более ёмкий сигнал. В сетях «Нового радио» (New Radio) 5G-NR

максимальная допустимая ширина полосы частот одного радиоканала увеличена рабочей группой LTE по сравнению с сетями 4G-LTE с 20 МГц до 100 МГц [2], как представлено в табл. 3.

Таблица 3

Количество РБ в полосе частот FR1- 5G в зависимости от  $\Delta F$

Разнос $\Delta F$ между под- несущими, кГц	Ширина полосы сигнала, МГц											
	5	10	15	20	25	30	40	50	60	80	90	100
	$N_{РБ}$ – количество ресурсных блоков внутри полосы сигнала											
15 кГц	25	52	79	106	133	160	216	270	-	-	-	-
30 кГц	11	24	38	51	65	78	106	133	162	217	245	273
60 кГц	-	11	18	24	31	38	51	65	79	107	121	135

Рассчитаем, можно ли использовать сигналы сети **FR1-5G** «Нового радио» для нашей задачи. Рассмотрим сигнал наиболее ёмкий из табл. 3 с полосой 100 МГц, с разнесом по частоте  $\Delta F = 30$  кГц. Количество ресурсных блоков внутри общей полосы сигнала задано 273. Каждый РБ переносит 12 поднесущих. Каждая поднесущая переносит 7 символов. Общее количество символов в сигнале составляет  $273 \cdot 12 \cdot 7 = 22932$  символа. При использовании модуляции поднесущих 256QAM, каждый символ которой переносит 8 бит, сигнал переносит 183456 бит. Этого недостаточно для нашего пакета 400000 бит. В структуре сигнала FR1- 5G предусмотрено недостаточно ресурсных блоков для работы с тяжёлым трафиком. Рассмотрим возможность использования сети в полосе FR2-5G.

На рис. 2 изображена ресурсная сетка 5G-NR[2]. Здесь: ресурсный элемент (РЕ) – минимальная единица для переноса информации в сети радиодоступа 5G-NR, ресурсный блок (РБ) – совокупность ресурсных элементов (РЕ), формируемая двенадцатью поднесущими на длительности одного слота;  $N_{ПН}^{РБ} = 12$  – количество смежных поднесущих частот в одном ресурсном блоке (РБ);  $W_{РБ}^u$  – ширина полосы ресурсного блока (от 180 кГц до 2880 кГц в зависимости от  $u$  - нумерации РБ);  $N_{СИМВ}^{СЛОТ}$  - количество символов в одном слоте (12 – для нормального префикса и 14 – для расширенного);  $N_{СЛОТ}^{СУБКАДР}$  - количество слотов в одном субкадре (от 1 до 16 в зависимости от нумерации типа сигнала);  $N_{СЛОТ}^{КАДР}$  – количество слотов в одном кадре (от 10 до 160 в зависимости от нумерации);

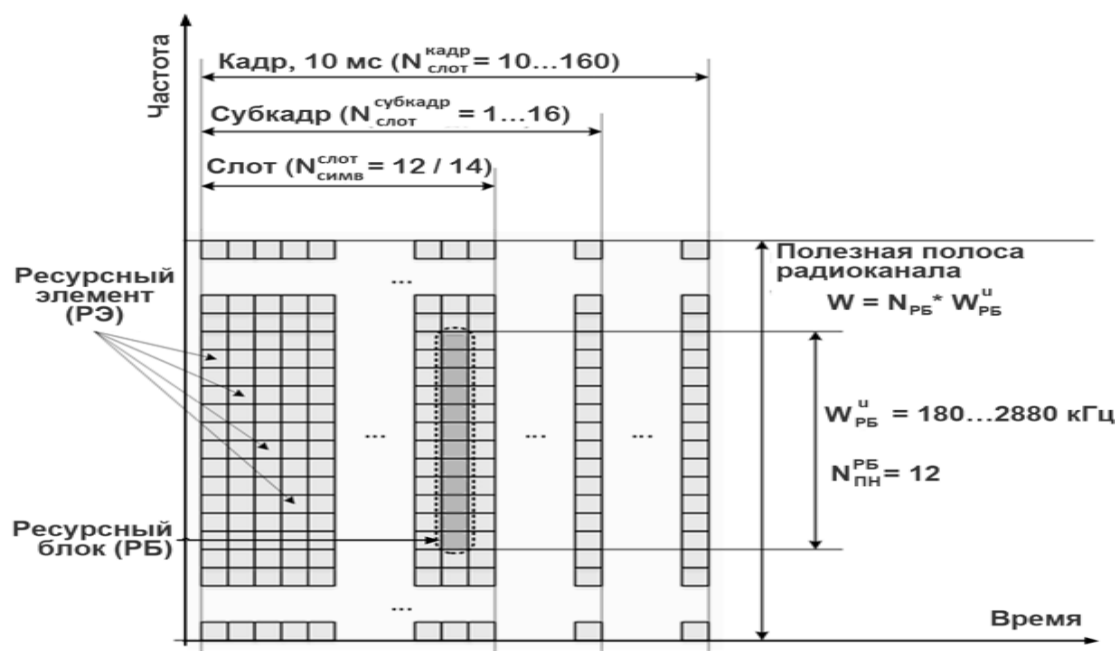


Рис.2. Ресурсная сетка 5G-NR

Рассмотрим общую универсальную структуру интерфейса новой сети 5G-NR, приведённую в последних релизах спецификаций 3GPP[2]

### Структура радио интерфейса сети 5G-NR

В отличие от сетей LTE-4G[3], где используется стандартная ширина спектра поднесущей 15 кГц, в новых пятого поколения сетях 5G-NR используются OFDM поднесущие с различной шириной спектра:  $\Delta F = 15$  кГц, 30 кГц, 60 кГц, 120 кГц, 240 кГц (см. табл. 4) и даже предусмотрена дополнительная полоса 480 кГц. Использование различных номеров типов РБ открывает широкие возможности для гибкой настройки поведения сети при предоставлении тех или иных услуг, например, для приложений критичных к уровню задержек (URLLC), уместно использовать ПН с широким спектром при меньшей длительности символа  $T_s$  и кадра, и наоборот, при передаче трафика широкополосного высокоскоростного доступа к Интернет (eMBB) и низкоскоростного трафика Интернета вещей (IoT) – целесообразно использовать "узкие" ПН. Но при этом необходимо точно рассчитать, какие объёмы пакетов может переносить тот или иной тип РБ, чтобы согласовать интенсивность входного потока с интенсивностью потока передач. Кратко рассмотрим основные понятия сети радиодоступа 5G-NR в контексте сравнения с сетями 4G-LTE, но самое главное – в контексте решения задачи синтеза видеосети тяжёлого трафика.

Таблица 4

Нумерация типов РБ

Номер типа РБ (u)	Ортогональный разнос по частоте поднесущих $\Delta F = 15 \cdot 2^u$ [кГц]	Циклический префикс (CP)
0	15 (соответствует LTE-4G)	нормальный
1	30	нормальный
2	60	нормальный, расширенный
3	120	нормальный
4	240	нормальный

### Кадровая структура технологии 5G-NR (кадр, субкадр, полуквадр, слот, символ)

- в 5G-NR передача данных в восходящем от мобильной станции (Uplink – UL) и нисходящем от базовой станции (Downlink – DL) направлениях организуется на основе кадров (frame) длительностью  $T_f = 10$  мс;
- каждый кадр делится на 10 субкадров длительностью  $T_{sf} = 1$  мс каждый;
- также каждый субкадр 1 мс делится на два полуквадра (полуквадр 0 и полуквадр 1) длиной по 0,5 мс;
- каждый субкадр делится на слоты (slot); количество которых в субкадре определяется шириной спектра поднесущей (т.е. нумерацией (u) типов ресурсных блоков (РБ)) и составляет 1, 2, 4, 8 или 16 слотов на субкадр (в отличие от сетей LTE-FDD с дуплексом с частотным разделением, где используется единая структура с двумя слотами на каждый субкадр при  $\Delta F = 15$  кГц).

Кадровая структура технологии 5G-NR [2] приведена в табл. 5 и на рис. 3 и рис.4.

Таблица 5

Кадровая структура в зависимости от полосы поднесущей (номера типа РБ)

Нумерация типов РБ (u)	u = 0	u = 1	u = 2	u = 3	u = 4
Ширина полосы поднесущей (ПН), кГц (ортогональный разнос между ПН)	15	30	60	120	240
Количество слотов в субкадре длиной 1 мс равно $2^u$	1	2	4	8	16



Количество слотов в кадре длиной 10 мс	10	20	40	80	160
Длительность слота, мс	1,0	0,5	0,25	0,125	0,0625
Длительность символа, мкс	66,7	33,3	16,6	8,33	4,17
Длительность обычного циклического префикса CP, мкс	4,7	2,41	1,205	0,6	0,3
Длительность расширенного CP, мкс	-	-	4,2	-	-



Рис.3. Слотовая и символьная структуры кадра в 5G-NR

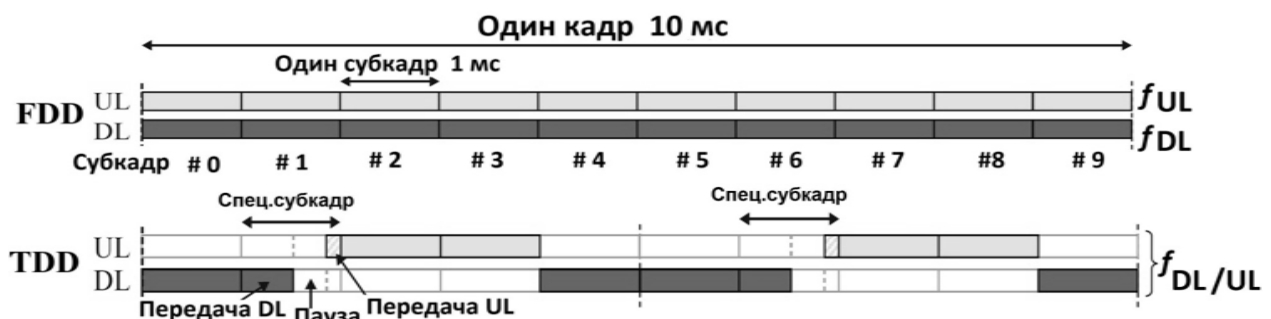


Рис. 4. Структура кадра с учётом вида дуплекса в 5G-NR

В 5G-NR распределение между направлениями передачи (UL/DL каналами) осуществляется внутри специального субкадра длительностью 1 мс на уровне символа, как показано на рис.4. Каждый OFDM символ слота может иметь тип "downlink" (D) (нисходящий канал), или "uplink" (U) (восходящий канал) или "flexible" (F) (гибкая конфигурация). В сетях 4G-LTE с временным дуплексом (TDD – Time Division Duplex) данное распределение направления передачи осуществляется внутри кадра длительностью 5 мс, на уровне субкадра длиной 625 мкс, состоящего из 6-ти символов. Один символ используется как пауза для смены направления передачи. В линии LTE-4G с временным дуплексом TDD разделение передач происходит внутри **слота** длительностью 0,5 мс, содержащего 6 или 7 символов. Пауза перехода направления передачи также находится в рамках одного символа порядка 70-80 мкс.

### OFDM символы в стандарте нового радио 5G-NR

В 5G-NR (New Radio) количество OFDM символов, приходящихся на один слот ( $N_{\text{СИМВ}}^{\text{слот}}$ ), не зависит от номера ( $u$ ) типа РБ и по аналогии с сетями 4G-LTE определяется только типом циклического префикса: 14 символов на слот для нормального префикса и 12 символов – для расширенного префикса. На рис. 5 приведена диаграмма OFDM-символов на сетке координат "Частота" - "Время". Длительность и ширина полосы символов связаны соотношением неопределённости (их произведение равно 1). Как видно из рис. 5, в новом поколении длительность символов уменьшена по сравнению с 4G ( $u = 0$ ).

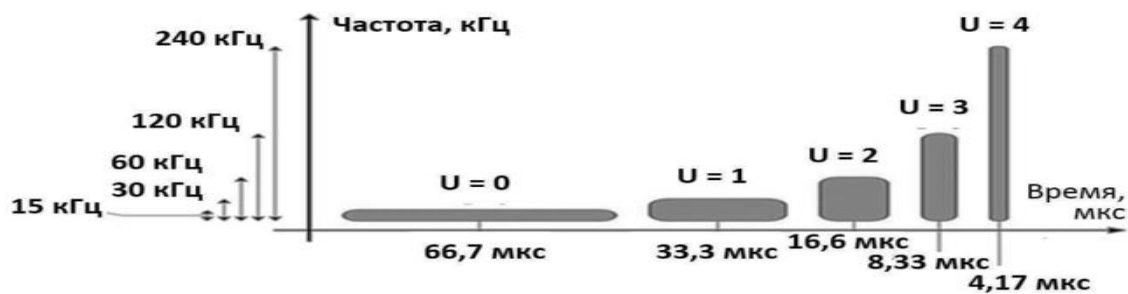


Рис. 5. OFDM символы 5G-NR

#### 1.4. Ресурсная сетка

По аналогии с 4G-LTE единицей частотно-временного ресурса в сетях 5G-NR является ресурсный блок РБ. Каждому абонентскому терминалу на определённый период времени (слот) для приема/передачи данных выделяется некоторое, определённое планировщиком системы, количество ресурсных блоков. Каждый ресурсный блок в частотной области содержит  $N_{\text{ПН}}^{\text{РБ}} = 12$  смежных поднесущих частот. Ширина полосы частотного ресурсного блока  $W_{\text{РБ}}^u$  является функцией номера его типа ( $u$ ) (см. Табл. 6).

Таблица 6

Ширина полосы частот одного ресурсного блока в зависимости от номера типа РБ

Номер типа РБ ( $u$ )	Ортогональный разнос частот $\Delta F \cdot 2^u$ [кГц]	Число поднесущих в РБ $N_{\text{ПН}}^{\text{РБ}}$	Ширина полосы частот ресурсного блока $W_{\text{РБ}}^u$ , кГц
0	15	12	180
1	30	12	360
2	60	12	720
3	120	12	1 440
4	240	12	2 880

#### Синтез протокола доступа к каналу большой видео сети высокого качества

Управление сетями на базе политики (PBNM – Policy Based Network Management) [4], является решающим фактором успешного выполнения современных быстрых операций и миссий. Для этого в зависимости от текущего сетевого сценария (количества узлов в сети) и заданной модели связанного трафика сначала синтезируется (или извлекается из памяти политики) оптимальный протокол множественного доступа к каналу, предварительно рассчитанный, моделированный и обслуживающий текущее количество узлов без коллизий случайного доступа и без помех от других станций и затем для этого протокола доступа синтезируется оптимальная структура мультиплекса сигнала OFDM+M-QAM с учётом отведённого частотно-временного ресурса, объёма пакета, и реального на момент передачи значения мощности сигнала и отношения «сигнал/шум» каждой несущей на входе приёмника.

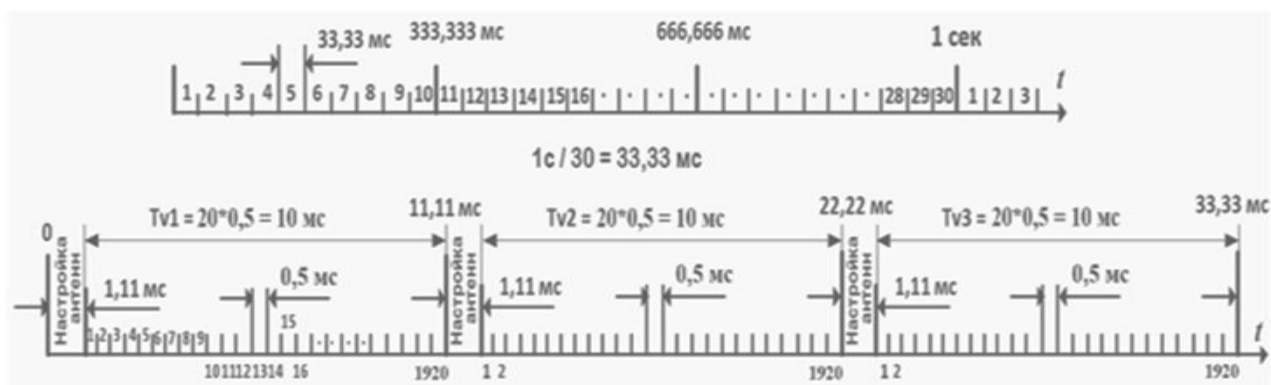
Исследуем процесс менеджмента сетью на базе политики (PBNM) на примере летающей видео сети (FANET) с изменяющимся сетевым сценарием (с изменяющимся числом узлов от 1 до  $N$ ) и заданной модели трафика. Начнём с модели трафика, когда каждый узел должен передавать по три видео потока – один от своей камеры и ещё два потока ретранслировать от двух ближайших соседей. Пакет видео потока имеет объём 50 кбайт (400000 бит = 0,4 Мбит) с учётом кодирования и сопровождающих данных и должен передаваться 30 раз в секунду. Скорость передачи данных для одного видео потока  $0,4 \times 30 = 12$  Мбит/с, для трёх потоков от одного узла - 36 Мбит/с. Требуемая пропускная способность каждого модема и сети в целом равна произведению 36 Мбит/с на количество  $N$  узлов в сети плюс 1/10 от этого произведения для учёта простоев передачи во время настроек антенн и приёмопередатчиков. Сначала для текущего сетевого сценария, например, 20 узлов в сети, и для заданной модели трафика необходимо синтезировать оптимальный протокол

множественного доступа к каналу, например, DTDMA – TDMA с направленными антеннами. Поскольку каждый корреспондент должен передавать в секунду по 30 пакетов каждого из трёх видеопотоков, то делим секунду на 90 равных частей по 11,11 мс. Это будет общая длительность цикла доступа к каналу. В начале каждого цикла отводим интервал 1,11 мс для настройки диаграмм направленности антенн и приёмопередатчиков на новые частоты при необходимости. На оставшемся интервале 10 мс кадра все 20 корреспондентов сети должны передавать последовательно по одному пакету за время слота  $10 \text{ мс}/20 = 0,5 \text{ мс}$ . Расчёт параметров протокола TDMA приведён на рис. 6.

$$\begin{aligned}
 N_{\text{узлов}} &:= 20 && \text{- количество узлов в равноранговой сети} \\
 N_{\text{бит}} &:= 400000 && \text{- объём пакета в битах} \\
 N &:= 30 && \text{- количество передач пакетов одного видео потока в секунду} \\
 N_{\text{потоков}} &:= 3 && \text{- количество видео потоков в секунду от одного узла} \\
 K &:= N_{\text{потоков}} \cdot N = 90 && \text{- количество пакетов в секунду от одного узла} \\
 T_{\text{цикла}} &:= \frac{1 \text{ с}}{K} = 11.111 \text{ мс} && \text{- период между передачами 1 пакета каждым узлом} \\
 T_{\text{антенны}} &:= 1.11111 \text{ мс} && \text{- время настройки антенны и р/с в начале каждого цикла} \\
 T_{\text{кадра}} &:= T_{\text{цикла}} - T_{\text{антенны}} = 10 \text{ мс} && \text{- общее время кадра множественного доступа к каналу для передачи одного пакета каждым узлом} \\
 T_{\text{слота}} &:= \frac{T_{\text{кадра}}}{N_{\text{узлов}}} = 500 \text{ мс} && \text{- длительность одного слота кадра доступа к каналу}
 \end{aligned}$$

**Рис. 6. Алгоритм расчёта оптимального протокола доступа к каналу TDMA**

Такая структура протокола доступа к каналу TDMA для 20 абонентов в сети приведена на рис. 7. Поскольку сигнал OFDM – это матрица распределения частотно-временных ресурсов, синхронизация по времени и частоте имеет решающее значение для достижения помехоустойчивости, близкой к потенциальной. Поэтому задача частотно-временной синхронизации решается активно и непрерывно передатчиком и приёмником каждого узла. Прогнозируя координаты приёмника и передатчика (высоту, широту, долготу) на момент передачи пакета, каждый узел сети должен вычислять расстояние трассы и задержку распространения сигнала, деля расстояние на скорость света ( $3 \cdot 10^8 \text{ м/с}$ ).



**Рис. 7. Протокол доступа к каналу TDMA для видео сети из 20 абонентов**

Эта оценка задержки распространения сигнала должна учитываться при формировании сигнала на передачу. Начало передачи должно опережать начало временного слота на время задержки, чтобы сигнал приходил на вход приёмника точно в начале слота. В результате время на задержку распространения сигнала в свободном пространстве не учитывается в каждом слоте синтезированного протокола TDMA. Задача обеспечения требуемой энергетики и надёжности каждой линии беспилотной сети решается с помощью

разработки и реализации плана траектории движения, который вводится в память каждого узла перед операцией и должен строго выполняться всеми узлами в ходе операции. В этом случае расстояния между узлами не превышают некоторого порога, гарантирующего требуемую энергетiku всех каналов. В противном случае, когда каждый узел двигается произвольно, получается хаос, и требуемой надёжности и достоверности связи никто не гарантирует. Например, задержка распространения на дальности прямой радиовидимости канала «Воздух-Земля» 370 км составляет 1,23 мс, это в 2,5 раза больше длины слота на рис. 6. Традиционный учёт этой задержки в виде защитного интервала внутри слота привёл бы к длительности слота 1,73 мс и уменьшению числа обслуживаемых узлов до 5 вместо 20 (в 4 раза). Кроме сокращения длительности слота доступа к каналу, знание дальности трассы позволяет оценить ослабление сигнала при распространении в свободном пространстве и прогнозировать мощность сигнала на входе приёмника, чтобы регулировать её до уровня, превышающего чувствительность приёмника для выбранной СКК. Если нет такой возможности, то выбирается другая СКК, имеющая меньшую чувствительность. Аналогичный метод используется для борьбы с Доплером с помощью предискажения передаваемого сигнала по частоте для компенсации доплеровского сдвига частоты. Такие методы компенсации Доплера и задержки распространения сигнала за счёт предискажений передаваемых сигналов реализуется в когнитивной авиационной системе связи «Воздух-Земля» гражданской авиации LDACS1 [7, 8].

После синтеза протокола TDMA, оптимального для заданных сетевого сценария и модели трафика по критериям минимальных задержки пакета и джиттера (рис. 6), следующим шагом разработки является синтез структуры мультиплекса сигнала (F-OFDM) (длительности его символа  $T_c$ , длительности префикса  $T_g$ , длительности полезной части сигнала  $T_u$ , разноса частот поднесущих  $\Delta F = 1/T_u$ , оптимальной по критерию спектральной эффективности (минимальной требуемой полосы частот для передачи пакета) для рассчитанного протокола TDMA. Речь идёт о выборе структуры сигнала OFDM, которая должна переносить видео пакет требуемого объёма (для заданной модели трафика) за заданное время слота, определённое синтезированным протоколом доступа к каналу) с максимально достижимой спектральной эффективностью (минимальной требуемой полосой частот) для текущей энергетике канала.

Исследуем, как количество символов OFDM сигнала на длительности слота влияет на спектральную эффективность сигналов OFDM. Для сравнительного анализа спектральной эффективности разных структур сигналов OFDM - с разным количеством символов на длительности слота (от 1 до 6) в табл.7 приведены результаты расчёта требуемых полос частот для передачи пакета объёмом 400000 бит в одном слоте для сетевого сценария из 20 узлов. Расчёты проведены для всех сетевых сценариев (числа узлов от 1 до 50).

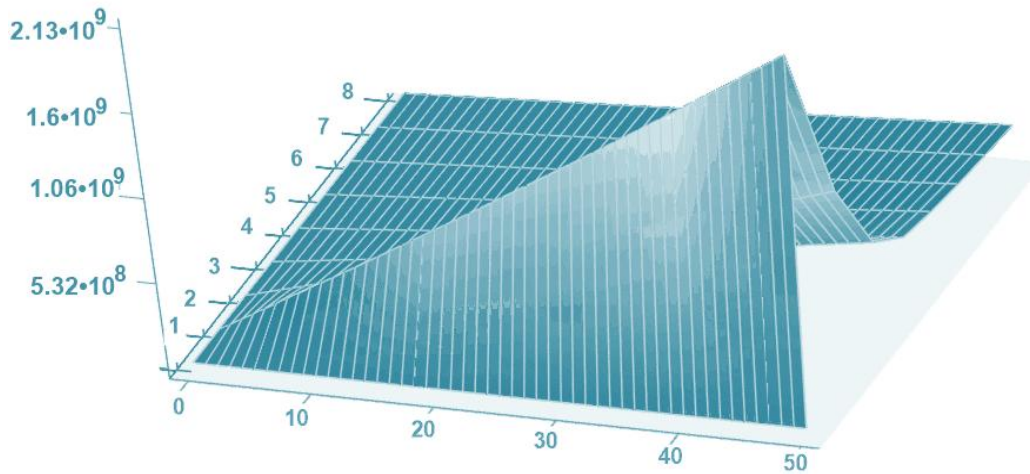
*Таблица 7*

**Сравнительная оценка спектральной эффективности сигналов OFDM с разным числом символов на слот для 20 узлов в сети**

Параметры	1 символ на слот	2 символа на слот	3 символа на слот	4 символа на слот	5 символов на слот	6 символов на слот
Длительность символа OFDM $T_s$ , мкс	500	250	166	125	100	83
Циклический префикс $T_g$ , мкс	12	12	12	12	12	12
Полезный интервал символа $T_u$ , мкс	488	238	154	113	88	71
Ортогональный разнос несущих $\Delta F = 1/T_u$ , кГц	2,049	4,2	6,49	8,85	11,36	14,08
Число символов в РБ $N_{\text{РБ}}^{\text{Симв}}$	12	24	36	48	60	72
Ширина полосы РБ $W_{\text{РБ}}$ , кГц	24,59	50,4	77,9	106,2	136,4	169
<b>Модуляция поднесущих 2PSK (1 символ переносит 1 бит)</b>						

Число бит, переносимых одним РБ	12	24	36	48	60	72
Требуемая полоса частот для передачи пакета 400000 бит, МГц	<b>820</b>	<b>840</b>	<b>865</b>	<b>885</b>	<b>909</b>	<b>940</b>
<b>Модуляция поднесущих QPSK (1 символ переносит 2 бита)</b>						
Число бит, переносимых одним РБ	24	48	72	96	120	144
Требуемая полоса частот для передачи пакета 400000 бит, МГц	<b>409</b>	<b>420</b>	<b>432</b>	<b>442</b>	<b>454</b>	<b>470</b>
<b>Модуляция поднесущих 16-QAM (1 символ переносит 4 бита)</b>						
Число бит, переносимых одним РБ	48	96	144	192	240	288
Требуемая полоса частот для передачи пакета 400000 бит, МГц	<b>205</b>	<b>210</b>	<b>217</b>	<b>221</b>	<b>227</b>	<b>234</b>
<b>Модуляция поднесущих 64-QAM (1 символ переносит 6 бит)</b>						
Число бит, переносимых одним РБ	72	144	216	288	360	432
Требуемая полоса частот для передачи пакета 400000 бит, МГц	<b>134</b>	<b>140</b>	<b>146</b>	<b>149</b>	<b>150</b>	<b>154</b>

Как видно из табл. 7, при модуляции поднесущих 2PSK структура сигнала OFDM с одним символом на слот требует ширины полосы сигнала **820 МГц** для передачи пакета объёмом 400000 бит за время 0,5 мс, что меньше **на 120 МГц**, чем требует структура сигнала OFDM с 6-ю символами на слот (**940 МГц**). При увеличении кратности модуляции MQAM поднесущих OFDM сигнала выигрыш в спектральной эффективности от уменьшения числа символов на длительности слота уменьшается, но он всегда имеет место для всех видов сигналов. Такие же расчёты требуемой полосы частот для передачи видео пакета за время одного слота 0,5 мс были сделаны для всех возможных сетевых сценариев (числа узлов от 1 до 50) и разных моделей трафика. Все расчёты показали более высокую спектральную эффективность сигналов с одним символом на слот, т.е. сигналов с самой большой длительностью символа, равной длине слота. И это естественно, поскольку мощность помех в полосе поднесущей уменьшается пропорционально увеличению  $T_c$  (уменьшению  $\Delta F$ ), а отношение с/ш при этом растёт. Поэтому мы выбираем для решения задачи построения видео сети тяжёлого трафика оптимальную по спектральной эффективности структуру сигнала OFDM с одним символом на слот протокола доступа, т.е. отказываемся от стандартного значения длины слота и разбиения его на 7 или на 6 символов, как в 4G-LTE и в 5G-NR. На рис.7 приведены результаты расчётов требуемой полосы частот ( $W_{TR}$ ) в МГц для разных сетевых сценариев (для количества узлов от 1 до 50), и для разных СКК (от 1 до 8) сигнала OFDM для передачи потоков при разных видах модуляции: 1 - 2PSK; 2- QPSK; 3 - 16QAM; 4 - 64QAM; 5 - 128QAM; 6-256QAM; 7-512QAM; 8 – 1024QAM для одной заданной модели трафика - видео пакета объёмом 400000 бит, который передаётся 30 раз с секунду в одном потоке, а каждый узел передаёт по три видео потока (по 90 пакетов в секунду). Из рисунка видно, что максимум требуемой полосы 2,13 ГГц имеет место, когда используется СКК 1 (2PSK) и все 50 узлов работают. Требуемые полосы частот при разном количестве узлов в сети (от 1 до  $N = 50$ ) были рассчитаны в Mathcad, и результаты представлены на рис.8 и в табл. 8.



**Рис.8. Зависимость требуемой полосы частот для передачи заданной модели трафика от числа узлов в сети (от 1 до 50) и от видов СКК (от 1 до 8)**

### **Адаптация сигнала OFDM по виду модуляции и кодирования поднесущих**

Вид сигнално-кодовой конструкции (СКК) должен выбираться автономно каждой мобильной станцией при условии, что прогнозируемая мощность принимаемого сигнала  $P_R$  на входе приёмника превышает уровень чувствительности приёмника  $P_{\text{чувст}}$  для данной СКК (PSK, MQAM) на 3-5 дБ, и, что прогнозируемое отношение «сигнал-шум» на входе приёмника превышает заданное значение для данной СКК для гарантии заданной достоверности приёма (вероятности ошибки на бит). Поэтому для многопараметрической адаптации радиолинии по СКК для каждого вида модуляции и кодовой скорости должны быть подготовлены таблицы зависимостей вероятностей ошибки на бит от отношения «сигнал/шум», вида схемы приёма и заданы чувствительности приёмника для разных СКК и схем приёма. Эти таблицы, создаваемые разработчиком модема по результатам тщательного моделирования работы модема в контролируемых условиях разных моделей каналов при реализации разных схем квазиоптимального приёма можно назвать таблицами политик многопараметрической адаптации радиолинии по СКК. В [8] приведены таблицы зависимостей  $P_{\text{ош}}$  от отношения «сигнал/шум» для разных СКК (разных видов модуляции и кодирования, и разных алгоритмов приёма – «жёсткого» поэлементного декодирования и «мягкого» декодирования «в целом» Витерби), узаконенных стандартами 802.11 и 4G (802.16m) для сигналов QPSK, 16QAM и 64QAM. Для реализации алгоритма многопараметрической адаптации по СКК радиостанция должна прогнозировать позиции передатчика и приёмника на момент передачи пакета, чтобы вычислить дальность трассы для прогнозирования потерь распространения сигнала и уровня мощности сигнала  $P_R$  и отношения сигнал/шум на входе приёмника, чтобы по ним выбрать оптимальный вид СКК с использованием таблиц политик СКК, для требуемой достоверности связи (заданного порога вероятности ошибки на бит). Процесс многопараметрической адаптации по видам СКК должен быть самообучаемым и реализуемым совместно с процессом оценивания физических параметров радиолинии. Алгоритм многопараметрической адаптации радиолинии по виду СКК приведён на рис. 9.

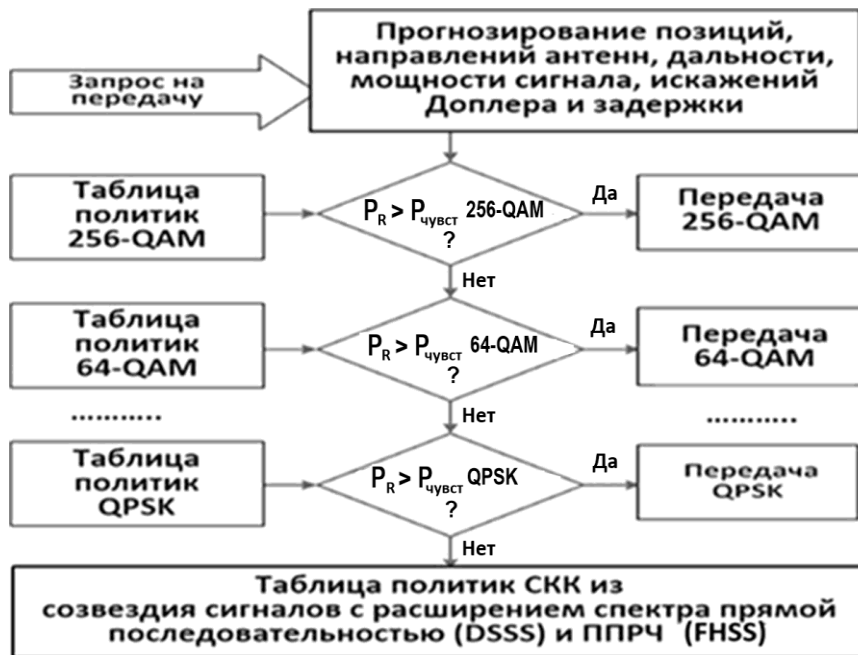


Рис. 9. Алгоритм многопараметрической адаптации радиолинии по виду СКК

Таким образом, процесс менеджмента сетью на базе политики (PBNM) представляет собой некий табличный режим управления сетью на базе многих таблиц политик (моделей), которые хранят знания результатов тщательного моделирования разных режимов работы сети и на базе реальных текущих измерений параметров канала и физического уровня. После выбора оптимальной СКК передатчик формирует в синхропреамбуле пакета информацию о выбранной СКК, необходимую для приёмника. Эту способность сигнала называют «автободом» – автоматическим определением при приёме вида сигнала и скорости передачи. Также в преамбуле указываются номера поднесущих, занятых помехами на приёмном и передающем концах, по которым данные не передаются. Пример таблиц политик для менеджмента сетью на базе политик (PBNM) приведён в табл. 8.

### Заключение

Исследованы существующие физические и каналные уровни стандартов мобильной связи четвёртого (LTE-4G[5], LTE-Advanced [3]) и пятого (NR-5G) [2] поколений мобильной связи с целью оценки их способности обеспечить функционирование видеосети большого размера с хорошим качеством изображения. Отрицательный результат этих исследований, свидетельствующий от том, что категория видео сетей тяжёлого трафика не входит в целевой профиль указанных стандартов, стал мотивацией для разработки новых физического и канального уровней, оптимизируемых совместно для большой мобильной видео сети высокого качества изображения.

Проведён сравнительный анализ спектральной эффективности сигналов OFDM с разным количеством символов на длительности одного слота, в котором передаётся пакет. Как показано в табл. 7, максимальную спектральную эффективность имеет сигнал с одним символом на слот. На базе этого факта был разработан алгоритм синтеза адаптивных протоколов физического и канального уровней, оптимизируемых совместно, который может быть использован для поддержки принятия решений в интеллектуальных мобильных самообучающихся видео сетях, задаваемых программой (SDN – Software Defined Network) и управляемых на базе политик (PBNM – Policy Based Network Management).

Таблица 8.

Таблица политик менеджмента сетью для одной модели трафика (90 пакетов от 1 узла в секунду) и разных (от 1 до 50) сетевых сценариев для 8-ми СКК

N Количество узлов в сети	СЭ макс б/с/Гц	СЭ мин б/с/Гц	C, Мб/с	Требуемая полоса частот (W, МГц) для разных N и СКК для одной модели трафика (пакет 400000 бит передаётся 30 раз в сек, по три потока от каждого узла)								Остаток полосы МГц	Остаток полосы МГц
	1024 QAM M=1024	BPSK M=2		M = 1024	1024 QAM M=1024	512 QAM M=512	256 QAM M=256	128 QAM M=128	64 QAM M=64	16 QAM M=16	QPSK M=4		
1	9.986	0.999	40	4.1	4.5	5.0	5.7	6.7	10.1	20	40	959	996
2	9.974	0.998	80	8.1	9.0	10.1	11.5	13.4	20.1	40	80.2	919	992
3	9.961	0.996	120	12.1	13.4	15.1	17.2	20.1	30.2	60.3	120.5	879	988
4	9.95	0.995	160	16.1	17.9	21.0	23.0	26.8	40.2	80.4	160.8	839	984
5	9.938	0.994	200	20.2	22.4	25.2	28.8	33.6	50.4	100.7	201.3	798	979
6	9.922	0.992	240	24.2	26.9	30.3	34.6	40.4	60.5	121.0	241.9	756	975
7	9.91	0.991	280	28.3	31.4	35.4	40.4	47.1	70.7	141.3	282.5	717	971
8	9.902	0.99	320	32.4	36.0	40.4	46.2	53.9	80.8	161.6	323.2	676	967
9	9.889	0.989	360	36.5	40.5	45.5	52.0	60.7	91.0	182.0	364.0	636	963
10	9.878	0.988	400	40.5	45.0	50.7	57.9	67.5	101.3	202.5	404.9	595	959
11	9.865	0.987	440	44.7	49.6	55.8	63.7	74.4	111.5	223.0	445.9	554	955
12	9.85	0.985	480	48.8	54.2	61.0	69.6	81.3	121.9	243.7	487.3	512	951
13	9.839	0.984	520	52.9	58.8	66.1	75.5	88.1	132.2	264.3	528.5	471	947
14	9.826	0.983	560	57.0	63.4	71.31	81.4	95.0	142.5	285.0	569.9	430	943
15	9.808	0.981	600	61.2	68.0	76.5	87.4	102.0	153.0	305.9	611.7	388	938
16	9.806	0.981	640	65.3	72.6	81.6	93.3	108.8	163.2	326.3	652.6	347	934
17	9.79	0.979	680	69.5	77.2	86.9	99.3	115.8	173.7	347.3	694.5	305	930
18	9.772	0.977	720	73.7	81.9	92.1	105.3	122.8	184.2	368.4	736.7	263	926
19	9.764	0.977	760	77.9	86.5	97.3	111.2	129.8	194.6	389.2	778.3	221	922
20	9.758	0.976	800	82.0	91.1	102.5	117.1	136.7	205.0	409.9	819.7	180	918
21	9.742	0.974	840	86.3	95.8	107.8	123.2	143.7	215.6	431.1	862.0	137	913
22	9.722	0.972	880	90.6	100.6	113.2	129.3	150.9	226.3	452.5	905.0	95	909
23	9.704	0.971	920	94.9	105.4	118.5	135.5	158.0	237.0	474.0	947.0	52	905
24	9.694	0.97	960	99.1	110.1	123.8	141.5	165.1	247.6	495.1	990.2	9	900
25	9.698	0.97	1000	103.2	114.6	128.9	147.3	171.9	257.8	515.5	1031	-31	896
26	9.67	0.967	1040	107.6	119.5	134.5	153.7	179.3	268.9	537.7	1075	-75	892
27	9.664	0.967	1080	111.8	124.2	139.7	159.7	186.3	279.4	558.7	1117	-117	888
28	9.658	0.966	1120	116.0	128.9	145.0	165.7	193.3	289.9	579.8	1159	-159	884
29	9.626	0.963	1160	120.6	133.9	150.7	172.2	200.9	301.3	602.5	1205	-205	879
30	9.628	0.963	1200	124.7	138.5	155.8	178.1	207.8	311.6	623.1	1246	-246	875
31	9.608	0.961	1240	128.1	143.4	161.4	184.4	215.1	322.7	645.2	1290	-290	870
32	9.598	0.96	1280	133.4	148.2	166.7	190.5	222.3	333.4	666.7	1333	-333	866
33	9.601	0.96	1320	137.5	152.8	171.9	196.4	229.2	343.7	687.3	1375	-375	862
34	9.586	0.959	1360	141.9	157.7	177.4	202.7	236.5	354.7	709.3	1418	-418	858
35	9.553	0.955	1400	146.6	162.9	183.2	209.4	244.3	366.4	732.7	1465	-465	853
36	9.538	0.954	1440	151.1	167.8	188.7	215.7	251.6	377.4	754.8	1509	-509	849
37	9.544	0.955	1480	155.1	172.3	193.9	221.5	258.5	387.7	775.3	1550	-550	844
38	9.536	0.954	1520	159.4	177.1	199.3	227.7	265.7	398.5	796.9	1594	-594	840
39	9.514	0.952	1560	164.0	182.2	205.0	234.2	273.3	409.9	819.7	1639	-639	836
40	9.518	0.952	1600	168.2	186.8	210.2	240.1	280.2	420.3	840.4	1681	-681	831
41	9.469	0.947	1640	173.2	192.5	216.5	247.4	288.7	433.0	865.9	1732	-732	826
42	9.49	0.949	1680	177.1	196.7	221.3	252.9	295.1	442.6	885.0	1770	-770	822
43	9.458	0.946	1720	181.9	202.1	227.3	259.8	303.1	454.6	909.2	1818	-818	818
44	9.458	0.946	1760	186.1	206.8	232.6	265.8	310.2	465.2	930.3	1861	-861	813
45	9.448	0.945	1800	190.5	211.7	238.2	272.2	317.5	476.3	952.4	1905	-905	809
46	9.428	0.943	1840	195.2	216.9	244.0	278.8	325.3	487.9	975.7	1951	-951	804
47	9.398	0.94	1880	200.1	222.3	251.0	285.8	333.4	500.1	1000	2000	-1000	799
48	9.406	0.941	1920	204.2	226.8	255.2	291.6	340.2	510.3	1020	2041	-1041	795
49	9.406	0.941	1960	208.4	231.6	260.5	297.7	347.3	520.9	1042	2083	-1083	791
50	9.398	0.94	2000	212.9	236.5	266.0	304.0	354.7	532.0	1064	2128	-1128	787



### Библиографический список

1. Рекомендация МСЭ-RM.2083-0 (09/1915 г.) - Концепция ИМТ–Основы и общие задачи будущего развития ИМТ на период до 2020 года и далее. 22 с.
2. 3GPP TS 38.104V17.5.0 (2022-03) 3rd Generation Partnership Project. Technical Specification Group Radio Access Network; NR; Base Station (BS) radio transmission and reception (Release 17).2022. 311 p. <http://www.3gpp.org>.
3. Recommendation ITU-R M.2012-3 (01/2018). Detailed specifications of the terrestrial radio interfaces of International Mobile Telecommunications-Advanced (IMT-Advanced).Electronic Publication. Geneva, 2018. 210 p.<http://www.itu.int/publ/R-REC/en>.)
4. **Y. Cai, Z. Qin, F. Cui, G. Y. Li, J. A. McCann** Modulation and Multiple Access for 5G Networks // Zhejiang University, Hangzhou, Imperial College London, Georgia Institute of Technology, Atlanta, USA. arXiv:1702.07673v1 [cs.OH] 21 Feb 2017. 31 p.
5. **Ермолаев, В.Т., Флакман, А.Г.** Теоретические основы обработки сигналов в беспроводных системах связи. Монография // Нижний Новгород: ННГУ, 2011. – 368 с.
6. **Perich, F.** Policy-Based Network Management for NeXt Generation Spectrum Access Control // Proceedings of IEEE. DySPAN, April 2007, 496-506pp.
7. L-DACS1 System Definition Proposal: Deliverable D3: Design Specifications for L-DACS1 Prototype // EUROCONTROL, Tech. Rep. Version 1.0. 2009.  
<http://www.eurocontrol.int/sesar-research/publications>
8. **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.

Е.В. Сидорова

ORCID: 0000-0002-6828-1625 e-mail: sidorova-elena@yandex.ru

К.А. Шишкина

ORCID: 0000-0002-2165-2193 e-mail: shishkinaksyusha03@gmail.com

Нижегородский государственный технический университет им. П.Е. Алексеева

Стремительное развитие в современном обществе информационных технологий послужило толчком к формированию электронных архивов документов в организациях, накопивших к настоящему моменту большое количество электронных документов. В работе проведен сравнительный анализ основных алгоритмов информационного поиска с целью выбора эффективного алгоритма для поиска текстовых документов в электронном архиве документов организации. Проведенный анализ включал ряд экспериментов, которые проводились с помощью программного продукта OpenServer на языке программирования PHP. Эксперименты основаны на сравнении методов поиска по отдельным критериям эффективности. Сравнение проводилось для алгоритмов двух групп методик: поиск в массиве данных и поиск подстроки в строке. Для проведения анализа были выбраны линейный, индексно-последовательный, бинарный и алгоритм поиска по дереву «Фибоначчи» из первой группы методов и алгоритм прямого поиска, алгоритм Рабина-Карпа, алгоритм Кнута-Морриса-Пратта и алгоритм Бойера-Мура из второй группы методов. В качестве критериев сравнения алгоритмов выбраны: релевантность, полнота и точность, уровень поискового шума, сложность и скорость алгоритмов. Полученные результаты проведенных экспериментов представлены в обобщенной таблице и визуализированы с помощью гистограмм Excel.

**Ключевые слова:** информационный поиск, критерии оценки эффективности методов информационного поиска, поиск в массиве записей, поиск подстроки в строке.

## ANALYSIS OF METHODS OF INFORMATION SEARCH FOR TAGS IN DOCUMENTS

E.V. Sidorova, K.A. Shishkina

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

**Purpose:** The goal of this article is to perform the comparative analysis of the main algorithms of the information search in order to choose the most effective algorithm for a search of text documents in a digital archive of the documents of the organization.

**Design/methodology/approach:** The conducted analysis includes a set of the experiments, which were performed using a program product OpenServer written with PHP program language. The experiments were based on comparison of the search methods divided by the efficiency criterion. The comparisons were conducted for two groups of the algorithms: a search in a data array and a search of the substring in a text string. The following algorithms were analyzed: a linear, an indexed sequential access method, binary, and a Fibonacci tree search as the first group and a direct search, the Rabin-Carp's algorithm, the Knuth-Morris-Pratt's algorithm, and Boyer-Moore's search algorithm as the second group. The next criteria were used for the comparison of the algorithms: a relevance, completeness and preciseness, a level of the search noise, and complexity and a speed.

**Findings:** The conducted analysis has showed that the Boyer-Moore's search algorithm exceeds other search methods in the relevance and the speed of the search.

**Research limitations/implications:** The Boyer-Moore's search algorithm will be used for a search of the documents in the digital archive of the target organization.

**Originality/value:** This research allows us to formulate a method of formatting the digital archive, which has been initially optimized for a search of the documents using the Boyer-Moore's algorithm, that in its turn leads to a decrease of the complexity and a level of the search noise and an increase of the speed and relevance of the search of the document in the archive.

**Keywords:** information search, effectiveness criteria of the search methods, search in a data array, search in a record array, search of a substring in the string.

Стремительное развитие в современном обществе информационных технологий послужило толчком к формированию электронных архивов документов в организациях, накопивших к настоящему моменту большое количество электронных документов, требующих хранения. Под электронным архивом будем понимать комплекс программных средств, которые предназначены для долгосрочного хранения в электронном формате проиндексированных документов. Проблему организации электронных архивов в организациях рассматривает Ивановский И.И. в своей работе «Особенности организации систем хранения электронных документов в архивах» [1].

Важным аспектом эффективной организации хранения электронных документов является осуществление быстрого и релевантного поиска информации в электронном архиве. Данная задача решается путем выбора и применения для поиска наиболее эффективных методов информационного поиска или их комбинации. На сегодняшний день разработано достаточно много методов поиска информации, в частности с использованием ключевых слов (тегов), и целью данной работы было проведение сравнительного анализа современных методов информационного поиска тегов в документах.

В научной литературе выделяют несколько определений «информационного поиска» [2]:

- «Информационный поиск» — это действия, методы и процедуры, позволяющие осуществлять отбор определенной информации из массива данных.
- Под информационным поиском понимается совокупность логических и технологических процедур, целью которых является выполнение запросов пользователей.
- Информационный поиск — это процесс поиска неструктурированной документальной информации, удовлетворяющий информационные потребности.
- Основной задачей информационного поиска является удовлетворение потребностей пользователей, которые формируются в виде запроса, представленного набором ключевых слов (тегов).

Среди методов информационного поиска для проведения экспериментов выберем два метода: поиск в массиве записей и поиск подстроки в строке (поиск по образцу).

Дадим краткую характеристику каждому из них.

**Поиск в массиве записей** представляет собой нахождение тех данных, которые удовлетворяют определенному свойству. Разновидностями поиска в массиве записей являются: линейный, индексно-последовательный, бинарный, интерполяционный и поиск по дереву «Фибоначчи» [3].

*Линейный (последовательный) поиск* — это простейший итеративный метод поиска заданного элемента на некотором множестве. Данный поиск осуществляется за счет последовательного сравнения последующего элемента с эталоном (ключом). Идея метода состоит в последовательном построении множества элементов, а так же в установлении направления просмотра элементов и получении результатов. Результатами поиска являются: нахождение или отсутствие индекса элемента. Линейный поиск осуществляется последовательно, на каждой итерации выполняются следующие шаги:

1) Задается одномерный массив элементов «А», размерности  $N$ . Задачей поиска является нахождение индекса  $k$  искомого заданного элемента.

2) Инициализируется значение переменной цикла  $i = 0$ .

3) Выполняется цикл, в теле которого происходит проверка ( $i \leq N$  и  $A[i]$ =искомому элементу), то элемент найден, запоминается индекс  $k=i$ , осуществляется выход из цикла. Иначе, элемент не найден, счетчик  $i$  инкрементируется, цикл повторяется до тех пор, пока не будет найден требуемый элемент или, когда просмотр массива завершится.

Достоинствами метода являются: простота реализации, отсутствие сортировки значений множества, работа в поисковом режиме, отсутствие выделения дополнительной памяти. Недостатком метода является низкая скорость перебора данных, которые имеют значительный объем.

*Индексно-последовательный поиск* — это метод поиска, который предназначен для повышения эффективности поиска в отсортированном файле. Идея метода: существует отсортированная таблица (файл). Дополнительно формируется еще одна таблица, которая имеет название «индекс». Элементы таблицы «индекс» состоят из ключа и указателя на запись в основной таблице.

Первым шагом выполнения поиска является проведение последовательного поиска в таблице «индекс». После проверки ключа, устанавливается нижняя граница поиска в основной таблице, а затем и верхняя. Более подробно метод изложен в [3].

Особенности индексно-последовательного поиска:

1. В индексированном последовательном поиске отсортированный индекс выделяется в дополнение к массиву.
2. Каждый элемент в индексе указывает на блок элементов в массиве или другой расширенный индекс.
3. Индекс ищется первым, затем в массиве и ведет поиск в массиве.

Достоинством метода является менее затратный по времени поиск, по сравнению с линейным. Недостатком является достаточный расход основной памяти.

*Бинарный (двоичный) поиск* — это метод поиска заданного элемента в отсортированном массиве, который происходит за счет последовательного сравнения, начиная с середины массива, следующего элемента с эталоном (ключом). Бинарный поиск осуществляется последовательным выполнением шагов, изложенных в [3].

Первым шагом метода является формирование упорядоченного по возрастанию массива элементов «А».

На втором шаге производится сравнение заданного элемента  $k$  с элементом, находящимся в середине массива.

На третьем шаге задаются первоначальные значения промежутка  $L = 1, R = N$ .  $L$  — индекс левого конца промежутка,  $R$  — индекс правого конца промежутка.

На четвертом шаге производится расчет индекса элемента с середины массива  $m = (L + R) \text{div} 2$ .

На пятом шаге сравниваем значение массива  $A[m]$  и заданного элемента. При выполнении сравнения возможно несколько результатов. Если значение заданного элемента больше значения массива, то процедура поиска в массиве не имеет значения и  $L = m + 1$ . Если значение заданного элемента меньше элемента массива, то при следующем осуществлении цикла можно считать  $R = m - 1$ . В случае равенства значений получаем результат — искомый элемент найден.

Процедура поиска завершается в случаях: нахождения элемента и когда индекс левого промежутка превышает индекс правого. Достоинствами метода являются: отсутствие выделения дополнительной памяти, быстрота выполнения алгоритма. Недостаток: предварительная сортировка массива.

*Интерполяционный поиск* — это метод поиска, основанный на понятии «интерполяция», т.е. осуществляется за счет процесса нахождения неизвестных значений на основе имеющихся. Данный метод поиска производится за счет деления массива на области и сужения этой области до одного элемента. Оценка выделенной области поиска производится по расстоянию между ключом поиска и значением элемента.

*Поиск по дереву «Фибоначчи»* — это усовершенствованный метод бинарного поиска, который основывается на концепции золотого сечения для выбора точки поиска в последовательности для процедуры поиска. Данный метод представляет собой АВЛ-дерево с минимальным количеством узлов при заданной высоте [3]. АВЛ — это аббревиатура, которая образована из первых букв создателей метода: Андельсона-Вельского Георгия Максимовича и Ландиса Евгения Михайловича. АВЛ-дерево — это иерархическая структура данных, которая содержит не больше двух потомков и имеет коэффициент балансировки  $k = 1$ . Коэффициентом балансировки называется некоторая характеристика, которая показывает значение отличия высоты левого и правого поддерева любого произвольного узла. Золотое сечение — это пропорциональное деление отрезка на неравные части, при котором весь отрезок так относится к большей части, как самая большая часть относится к меньшей, т.е. меньший отрезок так относится к большему, как большой ко всему [4].

Следующей группой методов является **поиск подстроки в строке (поиск по образцу)**, который представлен разновидностями алгоритмов: прямого поиска, Рабина-Карпа, Кнута-Морриса-Пратта, Бойера-Мура.

*Алгоритм прямого (последовательного) поиска* выполняет посимвольное сравнение строки и подстроки. Прямой поиск осуществляется последовательным выполнением шагов, изложенных в [5].

На первом шаге сравниваются первые символы строки и искомой подстроки.

Вторым шагом является процедура сравнения следующих символов строки и подстроки. Процедура завершается в том случае, когда подстрока найдена.

Переход к третьему шагу происходит если на предыдущем шаге не было совпадения символов строки и подстроки. На этом шаге сдвигается позиция подстроки в правом направлении и продолжается сравнение.

На четвертом шаге формируется условие останова алгоритма, т.е. достижение концов подстроки и строки.

Достоинствами алгоритма являются: небольшая сложность алгоритма –  $O(n)$  и распространенная документация для реализации метода. Недостаток: ограниченность применения для решения сложных задач.

*Алгоритм Рабина-Карпа* — это метод поиска строки за счет применения шаблона, т.е. применения процедуры хеширования. Идея алгоритма заключается в вычислении хеш-функции от образа вычислений и хеш-функции от части сопоставляемой исходной строки. Если данные хеш-функций равны, тогда происходит линейное сопоставление, в противном случае сопоставление не выполнено [5].

Под хешированием понимается процесс преобразования входного массива в выходную строку фиксированной длины.

Недостатками алгоритма являются: сложность  $O(n^2)$ , ограниченность применения для решения задач.

*Алгоритм Кнута-Морриса-Пратта* основывается на том, что после частичного совпадения начальной части образа в символьной строке мы имеем информацию, которая позволяет сдвинуться по строке больше, чем на единицу [6]. Алгоритм Кнута-Морриса-Пратта осуществляется последовательным выполнением следующих этапов.

На первом этапе происходит формирование таблицы  $d$  для осуществления процесса сдвига образа. Первый элемент этой таблицы равен минус единице. Остальные элементы вычисляются: значение  $d[j]$ , которое соответствует  $j$ -му символу образа и равен максимальному числу символов, которые предшествуют данному символу и совпадают с начальным образом. Если рассматриваемому символу предшествует  $k$  символов, то во внимание принимаются только  $k - 1$  предшествующих символов.

На втором этапе происходит сравнение символов образа строки, а так же вычисление сдвига образа в случае несовпадения. В случае несовпадения символов подстроки и строки образ сдвигается в правом направлении на величину:  $j - d[j]$ , где  $j$  – индекс текущего символа в подстроке,  $d[j]$  – значение таблицы этого символа.

Третий этап. Остановка алгоритма происходит в случае окончания строки или в случае успешного совпадения. [6]

Достоинства: применение для решения различного рода задач. Недостатки: сложность  $O(n^3)$ , ограниченность документации для реализации метода.

*Алгоритм Бойера-Мура*. Особенностью алгоритма является наличие некоторого количества предварительных вычислений над входной подстрокой. В этом случае, подстрока сравнивается с входным текстом не во всех позициях, так как часть сравнений, которые не приведут к результату, пропускаются [7]. Стоит обозначить понятие «образа», которое используется при выполнении этапов алгоритма. Под образом понимается подстрока, поиск которой требуется осуществить в процессе выполнения алгоритма.

Этапы выполнения алгоритма Бойера-Мура

На первом этапе происходит формирование таблицы  $d$  для осуществления процесса сдвига образа. Всем элементам таблицы  $d$  присваивается значение, которое равно длине образа. Затем каждому элементу таблицы  $d$  присваиваются значения, которые равны удаленности соответствующего символа образа от конца образа. В том случае, когда образ содержит несколько одинаковых символов, то элементу таблицы  $d$ , который соответствует данному символу, присваивается значение, которое равно удаленности от конца образа самого правого из одинаковых символов.

На втором этапе происходит поиск образа в строке. При сравнении образа со строкой, образ продвигается по строке слева направо. Однако, сравнения символов образа и строки выполняются, двигаясь по образу справа налево.

Третий этап. Сравнение образа происходит пока не будут достигнуты условия:

1. Найдено соответствие образа и строки.

2. Окончание строки.
3. Несовпадение символов образа и строки, что вызывает сдвиг образа вправо и процесс поиска повторится.

Достоинства: применение для решения различного рода задач, наличие распространенной документации для реализации метода. Недостатком является сложность  $O(n^3)$ .

Для проведения сравнительного анализа были выбраны линейный, индексно-последовательный, бинарный и поиск по дереву «Фибоначчи» из первой группы методов и метод прямого поиска, метод Рабина-Карпа, метод Кнута-Морриса-Пратта и метод Бойера-Мура из второй группы методов.

В качестве критериев оценки методов информационного поиска были выбраны: полнота поиска и точность поиска, уровень поискового шума, сложность и скорость алгоритмов [8].

Полнота поиска представляет собой количественную меру оценки, которая характеризуется способностью выдачи релевантных документов.

Точность поиска — это количественный показатель, позволяющий оценивать качество поиска.

Уровень поискового шума — это показатель, отражающий совокупность нерелевантных документов.

Для расчета полноты, точности уровня поискового шума применяются следующие обозначения:

1. Количество релевантных документов, выданных в результате запроса – А.
2. Количество нерелевантных документов, выданных в результате запроса – В.
3. Количество релевантных документов в поисковом массиве, не выданных в результате выполнения запроса – С.
4. Количество невыданных релевантных документов – D.

Расчет полноты поиска осуществлялся по формуле:

$$R = \frac{A}{A + C}$$

Расчет точности поиска осуществлялся по формуле:

$$P = \frac{A}{A + B}$$

Уровень поискового шума рассчитывался по формуле:

$$\text{ПШ} = \frac{B}{A + B + C + D}$$

Под сложностью алгоритма понималась характеристика оценки трудоемкости алгоритма с учетом затрат по времени его выполнения.

Временная сложность алгоритма представляет собой характеристику оценки трудоемкости алгоритма с учетом временных затрат. Она определяется как функция, которая зависит от длины строки. Временная сложность обозначается буквой «омикрон» —  $O(n)$ , где  $n$  — количество входных данных.

Разновидностями функции  $O(n)$  являются:

1. Константная сложность  $O(1)$  определяет сложность алгоритмов, выполнение которых не зависит от размера данных.
2. Линейная сложность  $O(n)$  определяет сложность алгоритмов, которые работают с одномерными массивами.
3. Полиномиальная сложность  $O(n^2)$ ,  $O(n^3)$  определяет сложность алгоритмов, которые работают с многомерными массивами.
4. Логарифмическая сложность  $O(\log(n))$  определяет сложность алгоритмов, которые работают для решения задач, связанных с декомпозицией.
5. Экспоненциальная сложность  $O(2^n)$  определяет сложность алгоритмов, которые применяют полный перебор.

Рассмотрев методы информационного поиска и определив критерии для оценки их эффективности, был проведен ряд экспериментов на тестовых наборах данных и выполнен сравнительный анализ полученных результатов.

Сравнительный анализ представляет собой метод оценки объектов, который происходит за счет сравнения одного объекта с другими по выбранным показателям. Выполнение сравнительного анализа происходило в несколько этапов.

На первом этапе выбраны объекты, подверженные сравнению, в нашем случае это 2 группы методов информационного поиска: поиск в массиве данных и поиск подстроки в строке.

На втором этапе выбирались показатели сравнения. Показателями сравнения являются критерии оценки эффективности методов информационного поиска, о которых говорилось ранее.

На третьем этапе сравнительного анализа определялось превосходства одного объекта над другим. В рамках третьего этапа анализа был осуществлен ряд экспериментов.

Эксперименты проводились на персональном компьютере со следующими характеристиками:

- Процессор — Intel (R) core (TM) i7-6500U CPU@ 2.60 GHz.
- Оперативная память — 8,00 GB.

Программная реализация методов поиска выполнена на языке программирования PHP. При проведении экспериментов предварительно созданы файлы формата – doc. Размер и объем файлов экспериментов варьируются. В качестве тестового набора данных для проведения экспериментов были взяты лирические стихотворения золотого и серебряного века русской поэзии. В процессе отдельных экспериментов количество файлов и строк в файлах изменялось. Для визуализации результатов экспериментов выполнено построение гистограмм Excel.

Программная реализация алгоритмов поиска представлена в таблице 1.

Таблица 1

Реализация методов поиска

Функция	Описание
«linearSearch»	Линейный поиск
«binaruSearch»	Бинарный
«indexSequentialSearch»	Индексно-последовательный поиск
«InsertedSearch»	Инвертированный поиск
«fibonace»	Поиска по «дереву Фибоначчи».
«strpos»	возвращает последнее вхождение подстроки в строке( прямой поиск)
«hash»	Генерирует хеш для алгоритма Рабина-Карпа
«RabinKarp»	Алгоритм поиска Рабина-Карпа
«kmp»	Алгоритм поиска Кнута-Морриса-Пратта
«Suffixes»	Формирование таблицы сдвигов (алгоритм Бойера-Мура)
«GoodCharacters»	Формирование правильного сдвига. Правильный сдвиг осуществляется тогда, когда строка совпала с последним символом (суффиксом) подстроки, а предшествующий символ подстроки не совпал с суффиксом
«BadCharacters»	Формирование плохого сдвига. Плохой сдвиг осуществляется тогда, когда строка не совпала с суффиксом подстроки.
«BoyerMoore»	Алгоритм поиска Бойера-Мура

**Эксперимент 1.** Оценка скорости алгоритмов информационного поиска в массиве документов.

Постановка задачи: оценить скорость выполнения методов, реализующих алгоритмы информационного поиска в массиве, в зависимости от объема массива документов.

Описание эксперимента: для определения зависимости времени выполнения алгоритмов от объема данных, массив документов постепенно увеличивался путем копирования исходной тестовой последовательности файлов.

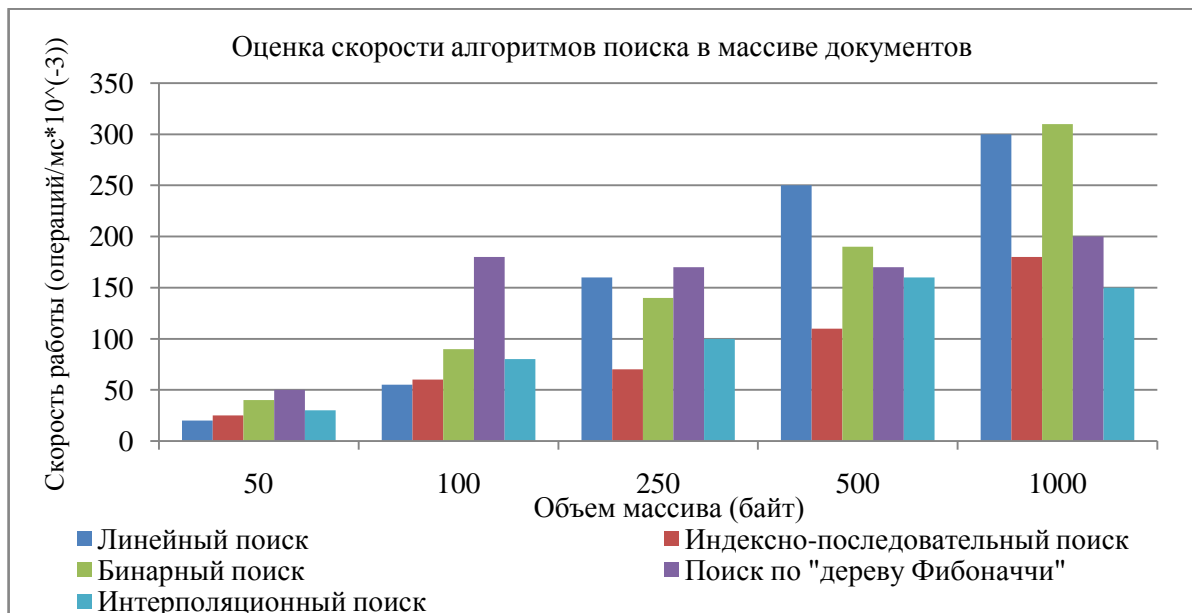
В качестве исходного массива были взяты лирические стихотворения: Пушкин А.С. «Я вас любил»; Тютчев Ф.И. «О как убийственно мы любим»; Брюсов В.Я. «Люблю одно»; Ахматова А.А. «Я не любви прошу»; Маяковский В.В. « Люблю»; Цветаева М.И. «Люблю, но мука еще жива»; Алексей Толстой «Моя душа летит приветом»; Пастернак Б.Л. «Душа»; Вадим Шефнер «Любовь минувших дней»; Гумилев Н.С. «Дремала душа».

Для оценки скорости использовалась встроенная функция `php microtime()` и подсчет количество итераций в тестируемых методах поиска.

В качестве тега для поиска выбрано слово «душа»

Эксперимент повторялся 100 раз на каждом объеме данных для получения несмещенной оценки скорости выполнения алгоритмов.

Результаты эксперимента представлены на рисунке 1.



**Рис.1. Гистограмма «Оценка скорости алгоритмов поиска в массиве данных»**

**Эксперимент 2.** Оценка скорости алгоритмов информационного поиска подстроки в строке.

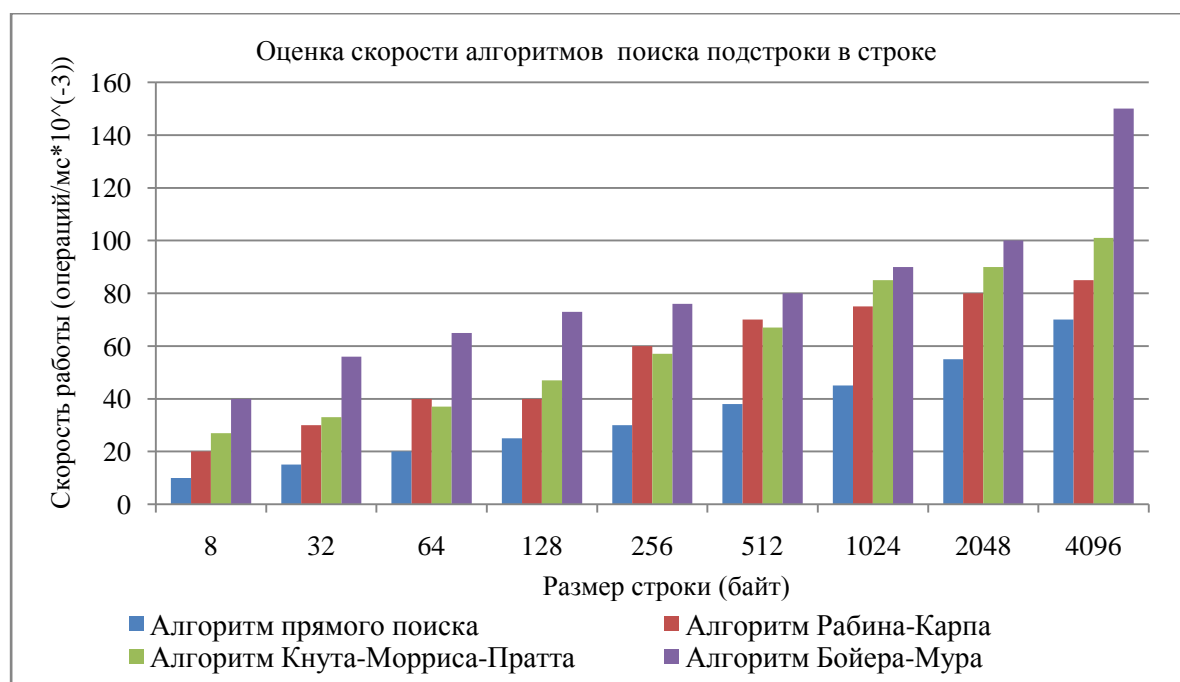
Постановка задачи: оценить зависимость времени выполнения методов, реализующих алгоритмы, от размера входной строки.

Описание эксперимента: для определения зависимости времени выполнения алгоритмов от размера входной строки на первом этапе использовались 10 тестовых документов, каждый из которых обрезался до объема всего в 8 символов. На втором этапе количество документов оставалось прежним, при этом объем каждого документа увеличился до 32 символов. Последующие этапы выполнялись по такому же принципу, т.е. увеличивался объем символов. В результате эксперимента объем символов устанавливался в следующие значения: 8, 32, 64, 128, 256, 512, 1024, 2048, 4096.

Для обрезки тестовых файлов до нужной длины использовалась функция `php truncate()`.

На диаграмме отображены средние значения скорости алгоритмов после выполнения серии экспериментов при зафиксированном размере строки.

Результаты эксперимента представлены на рисунке 2.



**Рис. 2. Гистограмма «Оценка скорости алгоритмов поиска подстроки в строке»**



### Эксперимент 3. Оценка сложности алгоритмов.

Постановка задачи: оценить зависимость между количеством операций, выполняемых алгоритмами, от количества входных данных для методов информационного поиска обеих группы.

Описание эксперимента: для определения зависимости между количеством операций от количества документов были выполнены шаги, на которых увеличивалось количество документов, аналогично 1 эксперименту. Сложность алгоритмов, записанная в легенде диаграммы через омикрон, взята из открытых источников. Результаты эксперимента представлены на рисунке 3.

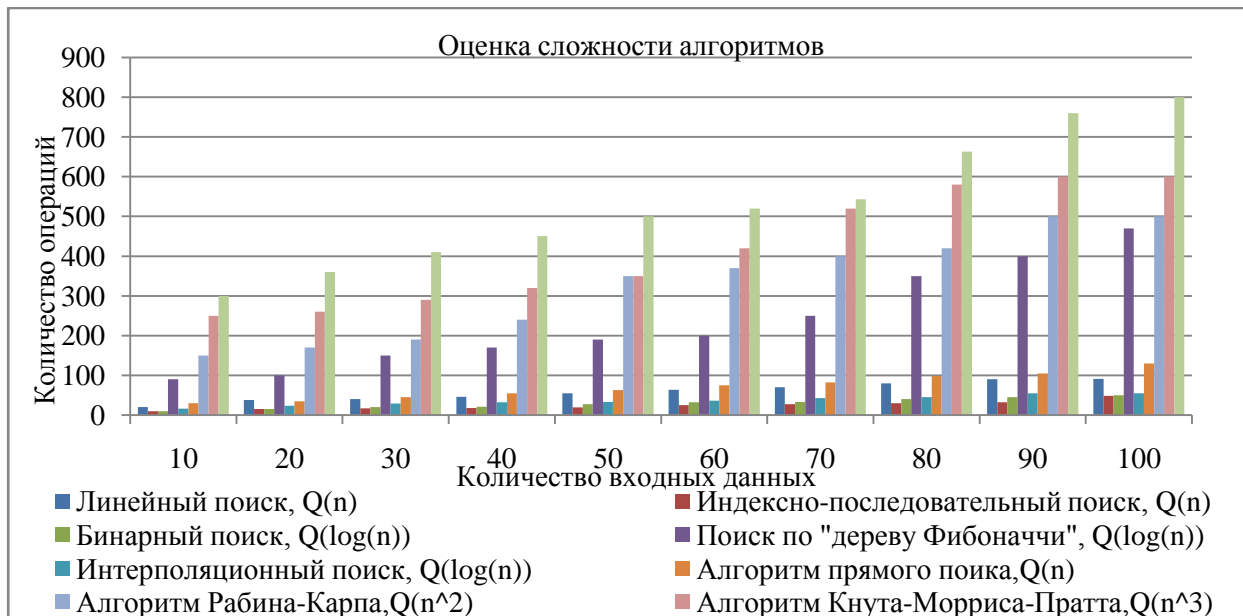


Рис. 3. Гистограмма «Оценка сложности алгоритмов»

### Эксперимент 4. Оценка релевантности.

Постановка задачи: выявить зависимость между найденными релевантными документами и количеством тегов в запросе, провести сравнение между алгоритмами обеих групп информационного поиска.

Релевантностью поиска называется степень соответствия запроса и найденного документа. Релевантность оценивалась при помощи следующих показателей: точность и полнота поиска, уровень поискового шума. Формулы для расчета этих показателей приведены ранее.

Описание эксперимента: для исходных тестовых документов, содержащих 10 лирических стихотворений, осуществлена последовательность запросов использованием указанных ранее функций рhr. При выполнении каждого шага количество тегов (ключевых слов) в запросе увеличивалось на 1. В эксперименте в каждый последующий запрос добавлялся новый тег. Первый запрос – тег «разлука». Второй запрос - теги «разлука любовь». Третий запрос - теги «разлука любовь душа». В результате выполнения функций по каждому алгоритму, были найдены документы, которые содержали указанные теги из запросов. Результаты расчета приведены в таблице 2.

Таблица 2

Расчет полноты, точности и уровня поискового шума методов поиска

Количество тегов в запросе	Точность поиска			Полнота поиска			Уровень поискового шума		
	1	2	3	1	2	3	1	2	3
Линейный поиск	0,8	0,75	0,6	0,88	0,85	0,7	0,17	0,2	0,25
Бинарный поиск	0,79	0,74	0,7	0,87	0,83	0,63	0,16	0,2	0,21
Интерполяционный поиск	0,79	0,76	0,64	0,86	0,84	0,68	0,17	0,17	0,2
Индексно-последовательный	0,8	0,7	0,6	0,85	0,8	0,75	0,13	0,22	0,25
Поиск по дереву «Фибоначчи»	0,74	0,73	0,67	0,88	0,78	0,67	0,12	0,2	0,22
Алгоритм прямого поиска	0,9	0,8	0,72	0,87	0,7	0,67	0,16	0,16	0,17
Алгоритм Рабина-Карпа	0,88	0,84	0,74	0,88	0,85	0,76	0,1	0,1	0,16
Алгоритм Кнута-Морриса-Пратта	0,88	0,84	0,74	0,89	0,87	0,72	0,1	0,12	0,16
Алгоритм Бойера-Мура	0,9	0,9	0,8	0,9	0,8	0,78	0,07	0,08	0,12

Наиболее релевантным алгоритмом поиска считается тот, у которого точность и полнота поиска имеют максимальные значения, а уровень поискового шума – минимальный.

**Подведем итоги.** Проведя ряд экспериментов в рамках сравнительного анализа, получили, что алгоритм информационного поиска Бойера-Мура превосходит другие методы поиска по скорости (эксперимент 3) и показателям релевантности информационного поиска (эксперимент 4). Именно поэтому алгоритм информационного поиска Бойера-Мура можно считать наиболее эффективным по сравнению с другими методами поиска. В дальнейшем он будет использован для организации поиска в создаваемом электронном архиве на целевом предприятии филиала АО «Корпорация Комета» - КБ «Квазар» в Нижнем Новгороде.

#### **Библиографический список**

1. **Ивановский, Н.И.** Особенности организации систем хранения электронных документов в архивах. – Москва: Вестник ВНИИДАД, ISSN:2610-1601, 2021. – 58-66с.
2. **Королев, О.Л., Бояджан, С.В.** Понятие информационного поиска. – Симферополь: ИП Бровко А.А., 2015. – 252с.
3. **Урвачева, В.А.** Обзор методов информационного поиска. – Ростов-на-Дону: Вестник таганрогского института им. А.П. Чехова, eISSN: 2306-2037, 2016. – 457-463с.
4. **Гренадеров, А.Н., Попов, А.П., Буданов, Б.В.** Золотое сечение. – Екатеринбург: Научный электронный журнал Меридиан, 2020. – 321-323с.
5. **Солдатова, Г.П., Татаринев, А.А., Болдырихин, Н.В.** Основные алгоритмы поиска подстроки в строке. – Ростов-на-Дону: ACADEMY, ISSN: 2412-8236, 2018. – 8-10с.
6. **Сухарев, А.И., Довгаль, В.М., Гордиенко, В.В.** Способ поиска по образцу. – Курск: AUDITORIUM, eISSN: 2312-4180, 2016. – 57-66с.
7. **Желудков, А.В., Макаров, Д.В., Фадеев, П.В.** Исследование алгоритмов точного поиска подстроки в строке. – Москва: Символ науки: международный научный журнал, ISSN: 2410-700X, 2016. – 61-62с.
8. **Касекеева, А.Б.** Исследование методов информационного поиска. – Астана: Big Data And Advanced Analytics, 2019. – 334-400с.

**ПРОЦЕДУРЫ ПОДДЕРЖКИ ПРИНЯТИЯ РЕШЕНИЙ ПРИ РЕАЛИЗАЦИИ  
ПАКЕТА ЗАДАНИЙ В РАСПРЕДЕЛЁННЫХ КОМАНДАХ, ОСНОВАННЫЕ  
НА УПРАВЛЕНИИ ПО КОМПЕТЕНЦИЯМ**

**Д.В. Баранов**

ORCID: 0000-0002-5902-6008 e-mail: smart.village.14@gmail.com

Нижегородский государственный технический университет им. Р.Е. Алексеева

В докладе предложен алгоритм принятия решений, основанный на обработке матриц компетенций, позволяющий использовать и сравнивать качество выполнения работ, а также сами компетенции, которые могут изменяться в течение времени. Исследование выполнено на тестовых выборках, сформированных при выполнении небольших проектов. Разработан алгоритм принятия решений при распределении пакета заданий. Проведено сопоставление предложенного подхода с существующими алгоритмами.

**Ключевые слова:** распределенные команды оценка компетенций, системы поддержки принятия решений, тестирование сотрудников.

**DECISION SUPPORT PROCEDURES FOR THE IMPLEMENTATION OF A  
PACKAGE OF TASKS IN DISTRIBUTED TEAMS ON THE BASIS  
COMPETENCE MANAGEMENT**

**D.V. Baranov**

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

**Purpose:** The author proposes an algorithm for making a decision on the distribution of a package of tasks under conditions of uncertainty, based on a combination of algorithms for processing competency matrices and probabilistic IRT models.

**Design / methodology / approach:** Is developed 3 steps of alghorythm:

1) The Rasch modelwith several employee competencies and job descriptions corresponding to the required (to complete the task) competenciesin virtual organisations.

2) Machine learning stage (model settings).

On the basis of the accumulated sample (characteristics of the task, the result of the implementation), the parameters of the model are evaluated - the competencies of the employee.

3) Application.

For a new task given by its characteristics, the probabilities of successful completion of the task by individual workers are found.

These probabilities are then used to solve the assignment problem.

**Findings:** The paper finds out optimal recruitment algorithm based on classical binary procedures for working with competency matrices, based onthe probabilistic models.

В статье найден оптимальный алгоритм подбора персонала на основе классических бинарных процедур работы с матрицами компетенций, оснащенных нечеткими вероятностными моделями.

**Research limitations/implications:** Research is limited to projects with a small number of performers. Nevertheless, the results obtained can be used in the future in project management systems.

**Originality/value:** The value of the entire article is originale algorithm that allows distribute tasks under conditions of uncertainty.

**Key words:** distributed teams, competency assessment, decision support systems, employment testing.

## **Введение**

Значимость темы подтверждается программой развития цифровой экономики России, где отмечен значительный объем использования современных информационно-коммуникационных систем [1].

Преимущества удаленной разработки – это экономия на ресурсах (нет постоянных социальных вычетов, больших затрат на аренду помещений, на передвижение), широкий выбор кандидатов по всему миру, рыночные условия ценообразования (фрилансеры устанавливают небольшую цену, зарабатывают за счёт объемов), необходимость социального дистанцирования и др.

Существующие меры не дают возможности выявить все риски переходного процессов подбора и оценки персонала, делая невозможной максимизацию эффективности предприятия [2]

## **Постановка задачи**

Цель работы – Создание, совершенствование, моделирование методов принятия решения о распределении пакета поручений в условиях неопределенности на основе комбинации алгоритмов обработки матрицы компетенций, вероятностных моделей современной теории тестирования (Item Response Theory, IRT) и самоорганизации системы управления [3, 4].

Для достижения поставленной цели необходима разработка:

- Модели компетенций сотрудников виртуальной организации;
- Модели вероятности правильного выполнения задания в виртуальной организации;
- Программных средства решения задачи распределения пакета заданий в условиях неопределенности.

В рамках статьи акцентируем внимание на задаче 1.

Основные концепции, используемые для решения этой задачи следующие.

Процедуры принятия решений совершенствуются через применение некоторых принципов построения мягких экспертных систем (сочетающих теорию нечетких систем, нейронные сети, вероятностные рассуждения, генетические алгоритмы и обладающая синергическим эффектом) [5]. Такие системы совмещают шаги вывода по нечетким продукциям с шагами многокритериального выбора решения.

Для управления персоналом предложено использовать управление по компетенциям – это подход, который сосредоточен на определении личностного потенциала сотрудников и направлен на стратегическое развитие компании, актуален при решении таких задач, как управление карьерой, перспективное обучение и развитие, оценка персонала, отбор персонала, организационные изменения, стимулирование персонала [6].

## **Существующие подходы и их развитие**

Опишем процедуры первичного подбора персонала, понятия матриц компетенций и др. Существуют различные технологии предварительной оценки сотрудников. Например, широко известна практика компаний Microsoft [7].

В этой компании процесс подбора начинается с прескрининга (то есть короткая первоначальной проверки кандидата [8]), после следует онлайн-тестирование со своими задачами. Процесс отбора кандидатов завершается очным тестированием.

В Microsoft и Amazon поведенческое собеседование интегрировано в технические сессии, потому что часто эти этапы проводят сотрудники из будущей команды. В тех

компаниях, где подход отличается, общение с будущими коллегами вынесено в отдельный этап. Например, Google определяет команду уже после успешного прохождения очного этапа. Самый оригинальный процесс у Facebook — выбор команды происходит уже после найма. Новый сотрудник 6 недель проводит в bootcamp — это тестовые проекты в разных командах, а затем принимается решение о найме.

Важным инструментом, используемым в рекрутинге и проектном управлении является матрица компетенций [2, 8].

**Таблица 1**

**Пример шаблона матрицы компетенций**

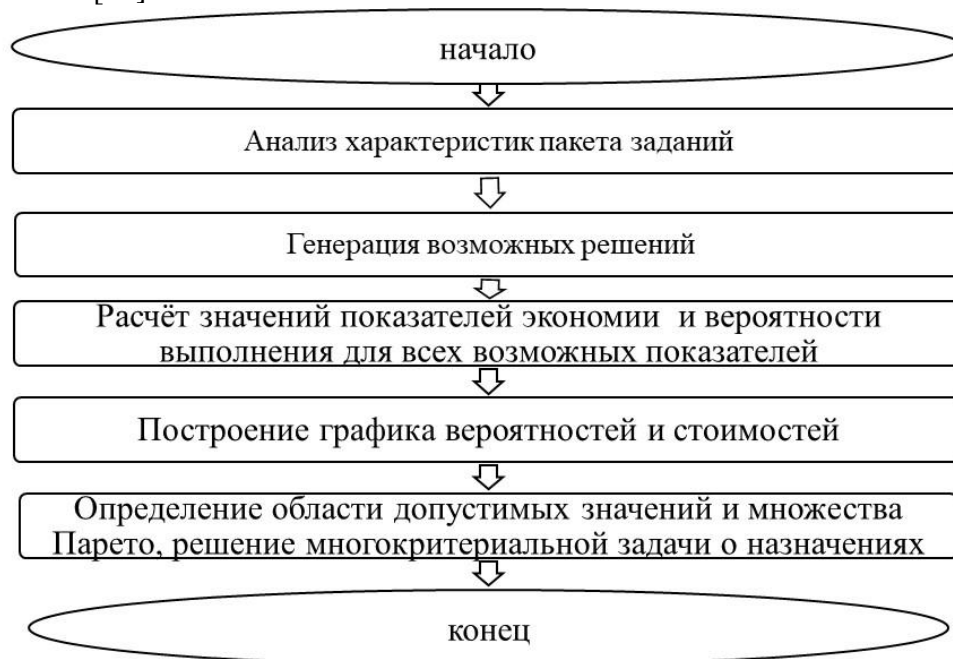
Проект	Роль	Компетенции				Срок
		1	2	...	N	
1	Задание 1	1. Ведущий программист				I кв
	Задание 2	1. Конструктор				
	Задание 3	1. Старший научный сотрудник				
2						II кв
3						

Затем за исполнение работ ставятся оценки и делаются выводы о дальнейших назначениях.

**Таблица 2 – Пример оценок за выполненные задания**

Сотрудник (ФИО)	Зад. 1	Зад. 2	Зад. 3	Зад. 4	Зад.5	Зад. 6	
Иванов А. Б.	5	4			4		
Петров С.В.	3	5		3		5	
Сидоров Г.Б.	4		2				
...							

На первом этапе работ (рис. 1) формируется матрица требуемых компетенций для персонала. Там компетенции расписываются по проектам и должностям, которые требуются для выполнения каждого проекта. Всем проектам соответствуют назначенные сроки. Предложен модернизированный алгоритм принятия решений при распределении пакета заданий в условиях неопределенности [4], система принятия решений также описана в статье [10].



**Рис. 1. Алгоритм принятия решений в условиях неопределённости**

Процедура оценки вероятности успешного выполнения задания может быть альтернативой традиционным оценкам компетенций сотрудников. По сравнению с тем, что есть в практике использования матриц компетенций, разрабатываемый подход предусматривает большую «гибкость» в принятии решений, что предоставляет руководителю значительную степень свободы (с бинарной логикой, можно никого не подобрать).

Для оценки вероятности успешного выполнения задания сотрудником предложено использовать модель Раша [11, 12]. Классификация этих моделей представлена на рис. 2.

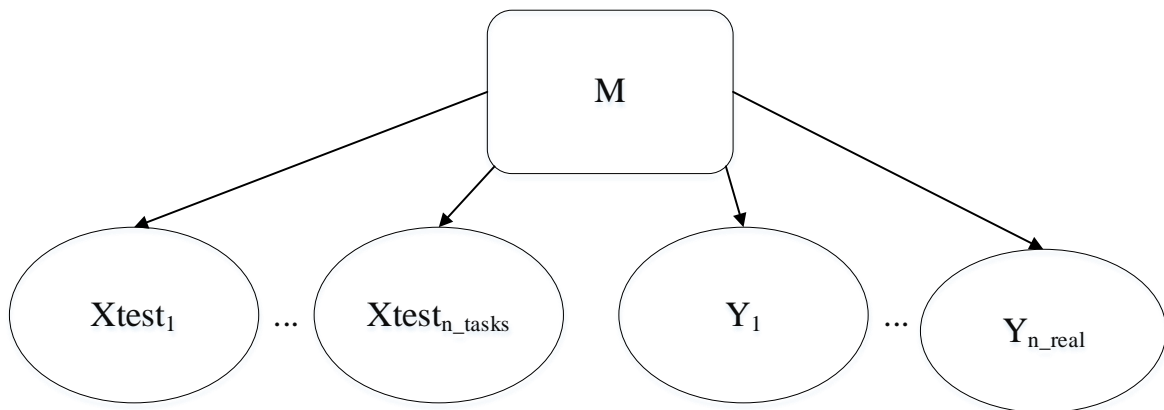
При вычислении рейтинга могут учитываться и рейтинги других участников киберсреды (так, рейтинг преподавателя высшего учебного заведения будет зависеть также от рейтингов его выпускников, у которых он является научным руководителем, а также от успеваемости студентов) [13].

Как вариант построения процедуры переоценки можно использовать Байесовскую сеть (БС) [14 – 16], основанную на следующих случайных величинах (Рис.3):

- Итоговую оценку обозначим  $M$ . Значения этой случайной величины: 1 – не выполнил работу, 2 – получил оценку «неудовлетворительно», 3,4,5 - это, соответственно, оценки «удовлетворительно», «хорошо», «отлично»;
- Для формирования оценки примем следующие показатели, две последовательности дискретных случайных величин:
  - $X_{test_i}, i=1, \dots, n\_tasks$  - оценки за тестовые задания
  - $Y_j, j = 1, \dots, m$ , - оценки за реальные задания



**Рис. 2. Классификация модели Раша**



**Рис. 3. Пример Байесовской сети**

### Выводы

Рассмотренные подходы применимы в таких областях, как:

- Системы управления проектами;
- Управление краудсорсинговыми командами;
- Системы управления компетенциями (образование).

### Библиографический список

1. Программа развития цифровой экономики в Российской Федерации до 2035 года // Электронный ресурс, режим доступа <http://static.government.ru/media/files/9gFM4FHj4PsB79I5v7yLVuPgu4bvR7M0.pdf>
2. Федорова, Л.А., Шиманский, А.А.. Разработка подходов к развитию человеческого капитала в корпоративной среде в условиях организации деятельности распределенных команд // Друкеровский вестник. 2021. № 3
3. Баранов, Д.В. Модели управления компетенциями в распределенных командах // Информ. Системы и технологии. ИСТ-2019: тез. докл. Междунар. науч.-техн. конф. НГТУ. – Н. Новгород, 2019. – с. 386 – 390

4. **Баранов, Д.В.** Процедура назначения заданий в распределенных командах с использованием модели Раша // Науч.-техн. вестник Поволжья. 2021. № 4. С. 17 – 24.
5. **Ярушкина, Н.Г., Наместников, А.М.** Мягкая экспертная система. Определение. Архитектура. Характеристики. Теоретические возможности и опыт разработки // Вестник УлГТУ, № 3 (200)
6. Управление по компетенциям - "must have" современной компании // Электронный ресурс, режим доступа <https://spb.hse.ru/dopbusiness/news/182104594.html#:~:text=%D0%A3%D0%BF%D1%80%D0%B0%D0%B2%D0%BB%D0%B5%D0%BD%D0%B8%D0%B5%20%D0%BF%D0%BE%20%D0%BA%D0%BE%D0%BC%D0%BF%D0%B5%D1%82%D0%B5%D0%BD%D1%86%D0%B8%D1%8F%D0%BC%D0%B8%20E2%80%93%20%D1%8D%D1%82%D0%BE%20%D0%BF%D0%BE%D0%B4%D1%85%D0%BE%D0%B4,%2C%20%D0%BE%D1%86%D0%B5%D0%BD%D0%BA%D0%B0%20%D0%BF%D0%B5%D1%80%D1%81%D0%BE%D0%BD%D0%B0%D0%BB%D0%B0%2C%20%D0%BE%D1%82%D0%B1%D0%BE%D1%80%20%D0%BF%D0%B5%D1%80%D1%81%D0%BE%D0%BD%D0%B0%D0%BB%D0%B0%2C>
7. Подготовка к собеседованиям в IT-гиганты: как я преодолела проклятье алгоритмического собеседования // Электронный ресурс, режим доступа <https://habr.com/ru/post/499394/>
8. **Просвирнина, Н.В.** Прескрининг как часть системы автоматизированного подбора персонала в компании // Московский экономический журнал № 4, 2020
9. **Купер, С.** Школа лидерства. Техники эффективного руководства. - СПб.: ИГ «Весь», 2014. – 224 с.
10. **Баранов, В.Г., Милов, В.Р., Суслов, Б.А.** Способы аналитической обработки для системы поддержки принятия решений // Информационно-измерительные и управляющие системы. 2010. № 2. С. 12-15.
11. **Баркалов, С.А., Моисеев, С.И., Кочерга, Н.С., Соловьева, Е.В.** Математические модели подготовки и проверки качества освоения компетенций в образовательном процессе / Открытое образование. – 2014. – № 2. – С. 9 – 16.
12. **Park, JY, Cornillie, F, van der Maas HLJ and Van Den Noortgate W A** Multidimensional IRT Approach for Dynamically Monitoring Ability Growth in Computerized Practice Environments. Front. Psychol. (2019) 10:620. doi: 10.3389/fpsyg.2019.00620.
13. Актуальные проблемы инфотелекоммуникаций в науке и образовании. IX Международная научно-техническая и научно-методическая конференция; сб. науч. ст. в 4 т. / Под. ред. С. В. Бачевского; сост. А. Г. Владыко, Е. А. Аникевич. СПб. : СПбГУТ, 2020. Т. 2. 748 с.
14. **Хлопотов, М.В.** Применение байесовской сети при построении моделей обучающихся для оценки уровня сформированности компетенций // Интернет-журнал «Науковедение», 2014 №5 (25) [Электронный ресурс] – М.: Науковедение, 2014 – 28 стр.- Режим доступа: <http://naukovedenie.ru/PDF/20TVN514.pdf>, свободный. – Загл. с экрана.
- Кустицкая, Т.А.** Прогнозирование успешности обучения студента с помощью байесовской сети. Информатизация образования и методика электронного обучения. Материалы III Международной научной конференции. В двух частях. Сибирский федеральный университет, Институт космических и информационных технологий. 2019. С. 257-262.
15. **Kustitskaya, T.A., Kytmanov, A.A., Noskov, M.V.** Student-at-risk detection by current learning performance indicators using bayesian networks. Journal of Functional Analysis. 2020. № 2004. С. 09774.



## ПОВЫШЕНИЕ ДОСТОВЕРНОСТИ МОБИЛЬНОЙ ЦИФРОВОЙ СВЯЗИ В НЕНАДЕЖНЫХ КАНАЛАХ

А.Д. Абакшин

ORCID: 0000-0002-0251-975X e-mail: abakshin.t@mail.ru

Нижегородский государственный технический университет им. Р.Е. Алексеева

Обеспечение безопасности передачи данных одна из важнейших задач современного общества. Это связано с интенсивностью развития различных систем цифровой связи, таких как, спутниковая, мобильная, космическая и др. Каждая из перечисленных систем использует для передачи беспроводные каналы, в которых на передаваемый сигнал действуют помехи различной физической природы. Одним из главных методов повышения надежности и эффективности обмена информации по каналам цифровой и спутниковой связи является помехоустойчивое кодирование. Особый интерес представляют корректирующие коды, ставшие доступными в следствии повышения вычислительных возможностей ЭВМ, в частности LDPC и полярные коды.

**Ключевые слова:** кодирование, полярные коды, LDPC, линейный код.

## INCREASING RELIABILITY OF MOBILE DIGITAL COMMUNICATION IN UNRELIABLE CHANNELS

A.D. Abakshin

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

**Purpose:** The purpose of this article is to improve the reliability of digital mobile communication due to the correction codes made available as a result of the increase in the power of electronic computers, that is, LDPC and polar codes.

**Design/methodology/approach:** The article discusses methods for protecting information from interference. Algorithms for encoding and decoding for polar codes and codes with a low density of parity checks are given. The experiments were based on a comparison of error-correcting coding methods according to the efficiency criterion. The comparison was made between LDPC code algorithms and between polar code algorithms. A comparison of the coding methods themselves was also carried out. The following coding algorithms have been described: construction of polar codes with the Arkan kernel, systematic coding for polar codes, LDPC code. The following decoding algorithms were analyzed: sequential elimination, list sequential elimination, stack sequential elimination, directed search, Sum-Product, min-sum, min-sum normalized. To compare the algorithms, we used the criterion of relativity of the level of bit errors relative to the noisiness of the communication channel.

**Findings:** Проведенный анализ показал текущую эффективность полярных и LDPC кодов. В ходе сравнения помехоустойчивых кодов установлено, что ни одного из приведённых в работе методов декодирования информации не позволяет полярным кодам превзойти LDPC в корректирующей способности при работе с большими блоками данных.

**Research limitations/implications:** the analysis will be carried out only on error-correcting codes used in digital mobile communications.

**Originality/value:** This study allows us to formulate a method for decoding data using polar or LDPC codes, which in turn will allow achieving high noise immunity of information if there is sufficient computing power to apply corrective algorithms.

**Keywords:** coding, polar codes, LDPC, line code.

## Введение

В последние годы происходит бурный рост телекоммуникационной отрасли, в связи с научно-техническим прогрессом, разрабатываются новые способы передачи информации, увеличиваются скорость и объем передачи данных. Прогресс несет множество положительных сторон, но не стоит забывать и об отрицательных. Повышению зашумленности каналов связи неразрывно связано с возникновением ошибок и помех в передаваемых данных. Логично, что повышения значимости помехоустойчивого кодирования из года в год только растет, как, собственно, и количество методов кодирования и декодирования информации. Особый интерес представляют коды, ставшие доступными в результате повышения мощностей вычислительной техники, коды с малой плотностью проверок на четность и полярные коды.

### 1. Полярные коды

*Поляризация* – это направление вектора электрической составляющей электромагнитной волны в пространстве.

В аспекте построения кода поляризация подразумевает эффект преобразования канала связи в  $N$  независимых его копий, в которых вероятность ошибки при передаче данных стремится к 0, или к 1. Виртуальные подканалы получаются за счет многократной рекурсивной конкатенации короткого кода ядра, который преобразует физический канал в виртуальные подканалы. Когда число рекурсий становится большим, виртуальные каналы, как правило, либо поляризуются, то есть обладают высокой надежностью, или становятся разреженными, не надежными. Каналы с единичной вероятностью ошибки называют замороженными.

#### 1.1 Полярные коды с ядром Арикана

Процесс построения полярных кодов с ядром Арикана описал П.К. Семенов[1] в своей работе «Практические аспекты алгоритма Тала-Варди списочного декодирования полярных кодов». Ядром Арикана называют матрицу  $G = \begin{pmatrix} 1 & 0 \\ 1 & 1 \end{pmatrix}$ , а через  $G^{\otimes s}$  обозначим  $s$ -ю кронекеровскую степень – матрицу размера  $2^s \times 2^s$ . Пусть  $B_{2^s}$  – это  $2^s \times 2^s$  – матрица такой перестановки для вектора длины  $2^s$ , что номер новой позиции  $i$ -го элемента получается как обратная запись представленного в виде  $s$ -битной последовательности числа  $i$ . Например,  $2 \equiv (1\ 0\ 0) \rightarrow (0\ 0\ 1) \equiv 1$ ,  $3 \equiv (0\ 1\ 1) \rightarrow (1\ 1\ 0) \equiv 6$ . Обозначим скалярный канал со входом  $u$  и выходом  $y$  как  $W(y|u)$ , отождествляя его с функцией плотности условной вероятности выходного символа. В случае передачи последовательности данных  $u_0^{N-1} = (u_0 u_1 \dots u_{N-1})$  размера  $N$  через канал  $W(y|u)$  будем рассматривать эквивалентный векторный канал  $W^N(y_0^{N-1}|u_0^{N-1}) = \prod_{i=0}^{N-1} W(y_i|u_i)$  с выходом  $y_0^{N-1} = (y_0 y_1 \dots y_{N-1})$ . В статье [2] предложен метод поляризации двоичного симметричного канала  $W^N(y_0^{N-1}|u_0^{N-1})$  без памяти с двоичным выходом, где  $N = 2^s$ , при передаче через него преобразованного блока данных  $u_0^{N-1} B_N G^{\otimes s}$ . Метод заключается в представлении  $W^N(y_0^{N-1}|u_0^{N-1})$  как совокупности  $\overline{W} \equiv \{W_N^{(i)}(y_0^{N-1}, u_0^{i-1}|u_i), 0 \leq i \leq N\}$  двоичных каналов, обладающими следующими свойствами:

1. Эквивалентность  $W$  и  $\overline{W}$ : суммарная пропускная способность  $\sum_{i=0}^{N-1} C(W_N^{(i)})$  равна  $C(W)$  при условии, что для каждого  $W_N^{(i)}(y_0^{N-1}, u_0^{i-1}|u_i)$  известны точные значения информационных символов  $u_0^{i-1}$  (на рис. 2.2 это условие обозначено прерывистой линией).

2. Асимптотическая поляризация каналов из  $\overline{W}$ : при  $N \rightarrow \infty$  относительная доля  $W_N^{(i)}$  в  $\overline{W}$  таких, что  $C(W_N^{(i)}) \rightarrow 1$ , стремится к  $C(W)$ , а пропускная способность остальных  $N(1 - C(W))$  каналов стремится к нулю.

Поляризационный эффект достигается применением невырожденного преобразования  $B_N G^{\otimes s}$  к данным. Таким образом, поляризующее преобразование канала  $W^{2N}(y_0^{2N-1}|u_0^{2N-1})$  основано на поляризации векторных каналов меньшей размерности  $W^N(y_0^{N-1}|u_0^{N-1}), i = N, \frac{N}{2}, \frac{N}{4}, \dots, 1$ .

В [2] показано, что эта декомпозиция для плотностей вероятностей  $W_N^{(i)}(y_0^{N-1}, u_0^{i-1}|u_i)$  выражается как

$$W_{2N}^{(2i)}(y_0^{2N-1}, u_0^{2i-1} | u_{2i}) = \frac{1}{2} \sum_{u_{2i+1}} W_N^{(i)}(y_0^{N-1}, u_{0,e}^{2i-1} \oplus u_{0,o}^{2i-1} | u_i \oplus u_{2i+1}) \times W_N^{(i)}(y_N^{2N-1}, u_{0,o}^{2i-1} | u_{2i+1}); \quad (1)$$

$$W_{2N}^{(2i+1)}(y_0^{2N-1}, u_0^{2i} | u_{2i+1}) = \frac{1}{2} W_N^{(i)}(y_0^{2N-1}, u_0^{2i-1} \oplus u_{0,o}^{2i-1} | u_{2i} \oplus u_{2i+1}) \times W_N^{(i)}(y_N^{2N-1}, u_{0,o}^{2i-1} | u_{2i+1}); \quad (2)$$

$$W_0^{(0)}(y|u) = W(y|u), \quad (3)$$

Где  $u_{0,e}^{2i-1}$  и  $u_{0,o}^{2i-1}$  обозначают компоненты вектора  $u_0^{2i-1}$  только с четными и нечетными номерами соответственно. Вероятности передаваемых битов в каналах  $W_N^{(i)}(y_0^{N-1}, u_0^{i-1} | u_i)$  оцениваются через рекуррентные формулы (1), (2) с условиями выхода из рекурсии (3). При  $N \rightarrow \infty$  канала  $W_N^{(i)}$  будет либо абсолютно без шума, либо абсолютно ненадежным. Поэтому информационные символы  $u_i$ , передаваемые по каналам плохого качества, можно считать всегда фиксированными («замороженными»). И если сопоставить  $i$ -ю строку матрицы  $G^{\otimes s}$  каналу  $W_N^{(i)}(y_N^{N-1}, u_0^{i-1} | u_i)$ , то  $r$  наиболее надежным каналам из  $\overline{W}$  соответствует линейный  $(2^s, r)$  – код длины  $2^s$  и размерности  $r$ , называемый полярным кодом. Таким образом, чтобы задать полярный код  $s$  достаточно указать множество  $A$  индексов неиспользуемых строк матрицы  $G^{\otimes s}$ .

## 1.2 Систематическое кодирование для полярных кодов

Идея алгоритма состоит в кодировании через декодирование:  $k$  символов кодового слова равны информационным символам, значения остальных  $n - k$  символов кодового слова полагаются стертными и восстанавливаются с помощью декодера. Алгоритм основан на следующих свойствах матрицы  $A^{\otimes m}$ , из  $k$  строк которой состоит порождающая матрица полярного кода:

1. Матрица  $A^{\otimes m}$ , связаны рекурсивным соотношением.

$$A^{\otimes(m+1)} = \begin{pmatrix} A^{\otimes m} & 0_{2m} \\ A^{\otimes m} & A^{\otimes m} \end{pmatrix}, \quad (4)$$

Где  $0_{2m} - 2^m \times 2^m$  нулевая матрица.

2.  $A^{\otimes m}$  является нижнетреугольной матрицей с единичной диагональю. Она всегда обратима.

3. Любая подматрица  $A_{D,D}^{\otimes m}$  матрицы  $A^{\otimes m}$ , состоящая из строк и столбцов с индексами из множества  $D \subset \{0, \dots, n-1\}$ , также является нижнетреугольной с единичной диагональю, и потому обратима.

При вычислениях используется рекурсивная структура порождающей матрицы, задаваемая формулой (4). Предполагается, что множество позиций информационных символов в кодовом слове совпадает с множеством  $N \subset \{0, \dots, n-1\} \setminus F$ ,  $|N| = k$ , где  $F$  - множество индексов замороженных символов во входной последовательности. При систематическом кодировании выполняется поиск решения системы линейных уравнений

$$c_0^{n-1} = u_0^{n-1} A^{\otimes m}. \quad (5)$$

В данной системе неизвестными являются элементы последовательностей  $c_{0,F}^{n-1}$  и  $u_{0,N}^{n-1}$ ,  $u_{a,D}^b$  - подпоследовательность последовательности  $u_a^b$ , состоящая из элементов  $u_i, i \in D$ . Благодаря рекурсивной структуре матрицы  $A^{\otimes m}$ , систему (5) можно представить в следующем виде.

$$c_{n/2}^{n-1} = c_{n/2}^{n-1} A^{\otimes(m-1)}, c_0^{n-1} - c_{\frac{n}{2}}^{n-1} = u_0^{\frac{n}{2}-1} A^{\otimes(m-1)}. \quad (6)$$

Таким образом, задача поиска решения системы (5) может быть сведена к поиску решений двух систем в вдвое меньшей размерности.

Алгоритм систематического кодирования для полярных кодов с ядром Арикана представлен и описан в работе [3] и продемонстрирован на рис. 1. На вход данному рекурсивному алгоритму подается кодовое слово  $c_0^{n-1}$  с заданными значениями информационных символов  $c_{0,N}^{n-1}$ , длина полярного кода  $n$ , множество замороженных символов  $F$ . Результатом работы алгоритма являются вычисленные значения проверочных символов  $c_{0,F}^{n-1}$ . На шагах 4 и 7 осуществляется переход к задаче систематического кодирования меньшей размерности.

Сложность алгоритма систематического кодирования Арикана можно выразить как  $N_n = \frac{1}{2} n \log n$  операций сложения для элементов  $F_2$  [3].

```

POLARSYSENCODE( $c_0^{n-1}, n, \mathcal{F}$ )
1  if ( $n = 1$ )
2  then if  $0 \in \mathcal{F}$ 
3      then  $c_0 \leftarrow 0$ 
4  else POLARSYSENCODE( $c_{n/2}^{n-1}, n/2, \{i - n/2 | i \in \mathcal{F}, i \geq n/2\}$ )
5       $\mathcal{N} \leftarrow \{0, \dots, n-1\} \setminus \mathcal{F}$ 
6       $v_{0,\mathcal{N}}^{n/2-1} \leftarrow c_{0,\mathcal{N}}^{n/2-1} - c_{n/2,\mathcal{N}}^{n-1}$ 
7      POLARSYSENCODE( $v_{0,\mathcal{N}}^{n/2-1}, n/2, \{i | i \in \mathcal{F}, i < n/2\}$ )
8       $c_{0,\mathcal{F}}^{n/2-1} \leftarrow v_{0,\mathcal{N}}^{n/2-1} + c_{n/2,\mathcal{F}}^{n-1}$ 

```

Рис. 1. Алгоритм Арикана для систематического кодирования полярных кодов

### 1.3 Алгоритмом последовательного исключения

Алгоритмом последовательного исключения (SC) предназначен для декодирования полярных кодов. Основан на расчете коэффициента правдоподобия для каждого символа на каждом уровне поляризации канала. Число узлов определяется как  $N(\log_2 N + 1)$ . Так же стоит отметить, что декодирование алгоритмом последовательного исключения является алгоритмом с мягким решением. Сам алгоритм представлен в [2].

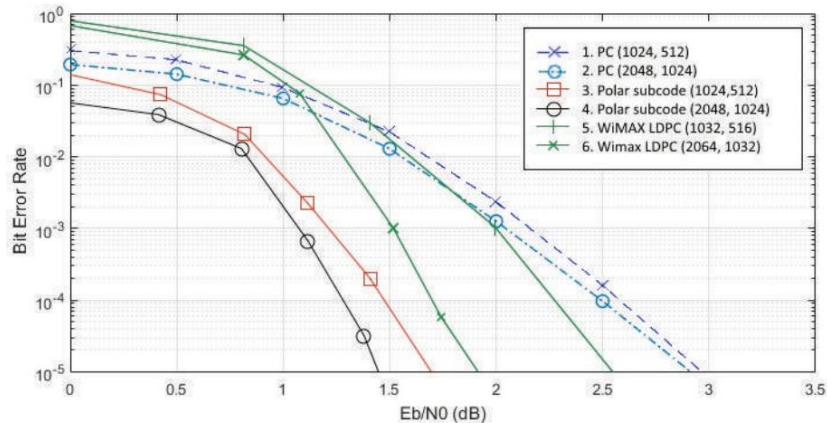


Рис 2. Показатель битовых ошибок для полярных код с декодером на основе алгоритма последовательного исключения, полярных подкодов и LDPC кодов

В исследованиях Тимофеева Г.С. приведены результаты моделирования результатов представлены на рис. 2. [4]: систематических полярных кодов с SC-декодированием (1024, 512) и (2048,1024) (кривые 1 и 2); полярные подкоды (1024, 512) и (2048,1024) (кривые 3 и 4) в соответствии с [5]; LDPC коды (1032,516) и (2064,1032) в соответствии со стандартом IEEE 802.16 Wireless MAN [6].

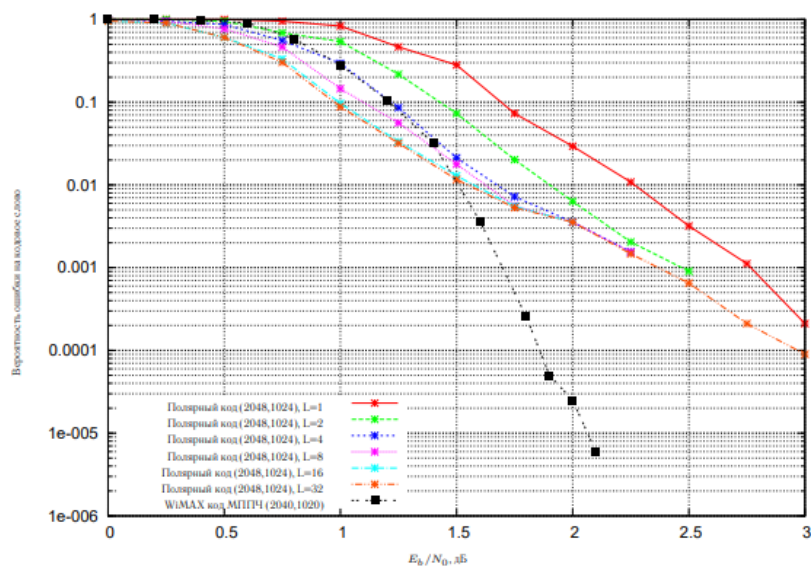
Видно, что при  $N \approx 1000$  полярные коды показывают сходный уровень битовых ошибок (BER) с LDPC кодами при показателях сигнал/шум  $< 2$  дБ. В остальных случаях полярные коды с SC-декодером проигрывают LDPC кодам, кроме уровня  $BER = 10^{-5}$ .

### Списочный алгоритмом последовательного исключения

Идея традиционного декодера [2] для полярных кодов заключается в последовательной оценке значений передаваемых битов по подканалам  $W_N^{(i)}(y_0^{N-1}, u_0^{i-1} | u_i)$ . При этом для (i+1)-го канала в качестве достоверных битов  $u_0^i$  используются оценки входов подканалов всех предшествующих номеров. Таким образом, если на какой-то стадии происходит ошибка, то и оценка всех оставшихся битов будет ошибочной. Алгоритм Тала-Варди позволяет исправить эту ситуацию.

В своих исследованиях Миславская В.Д. приводит график зависимости вероятности ошибки декодирования от отношения сигнал/шум для (2048, 1024) полярного кода и (2040, 1020) кода LDPC при декодировании которых использовались списочный алгоритм последовательного

исключения и алгоритм распространения доверия, соответственно [3]. График представлен на рис. 3.



**Рис 3. Вероятность ошибки декодирования (2048, 1024) полярного кода**

Можно увидеть, что при увеличении размера списка L вероятность ошибки декодирования полярного кода снижается. Случай  $L = 1$  соответствует классическому алгоритму последовательного исключения, в то время как случаи  $L = 16$  или  $L = 32$  соответствуют декодированию полярных кодов почти по максимуму правдоподобия, однако, их корректирующая способность оказывается существенно ниже чем кодов LDPC[3].

### Стековый алгоритм последовательного исключения

Отличие стекового от списочного алгоритма состоит в том, что в стеке (приоритетной очереди) хранятся пути различных длин. На каждой итерации декодирования из стека выбирается путь  $u_0^{i-1}$  с наибольшей вероятностью  $P(u_0^i | y_0^{n-1})$ . Этот путь удлиняется на элемент  $u_i$ , для  $u_0^i$  вычисляется вероятность  $P(u_0^i | y_0^{n-1})$ , затем путь  $u_0^i$  помещается в стек вместе с вычисленной вероятностью. Как и в случае списочного алгоритма, если  $i \in F$ ,  $F$  - множество замороженных символов, то  $u_i = 0$ , иначе рассматриваются оба варианта значения  $u_i \in \{0,1\}$ . Если число путей в стеке превосходит заданное ограничение, то пути с наименьшими вероятностями  $P(u_0^i | y_0^{n-1})$  исключаются из стека [3].

### Декодирование методом направленного поиска

Декодирование полярного кода по максимуму апостериорной вероятности предполагает поиск последовательности  $u_0^{i-1}$ , при которой достигается максимум  $P(u_0^i | y_0^{n-1})$  с учетом наличия ограничений на значения замороженных символов [3]. Предлагаемый подход берет свое начало из стекового алгоритма последовательного исключения, и обеспечивает существенное снижение сложности по сравнению со своим прообразом, при той же корректирующей способности.

## 2. LDPC коды

LDPC-код, как и любой линейный блочный код описывается проверочной матрицей, но на практике в основном используют двудольные графы Танстола. Они содержат в себе информацию о парах индексов строк и столбцов, на пересечении которых есть единица, однако могут использоваться и специальные способы задания матриц.

Процесс кодирования LDPC кода мало чем отличается от такового у кода Хэмминга, а именно разреженностью проверочной матрицы, то есть нулей в ней должно быть значительно больше, чем чего-либо другого.

Сегодня используется два основных метода задания матриц. В первом случае используют псевдослучайный генератор. Во втором используют специальные техники, основанные, например, на группах и конечных полях.

Для LDPCкодов существует два способа декодирования: жесткий и мягкий. «Жесткое» декодирование используется при небольшом количестве ошибок в канале, является самой простой, но далеко не самой оптимальной схемой. Мягкое декодирование куда более интересно в связи со своей эффективностью и сложностью реализации.

### 2.1 Алгоритм «Sum-Product»

Сообщение в процессе декодирования представлено вероятностью каждого полученного входного бита. Вероятности входных битов априорные, то есть известны еще до начала декодирования LDPCкода. Декодер возвращает апостериорные вероятности – логарифмическое отношение правдоподобия (LLR). LLRудобно тем, что его использование снижает вычислительную сложность дальнейшей работы с информацией. Целью алгоритма является вычисление максимальной апостериорной вероятности для каждого из входных закодированных битов[7]. Информация, полученная за счет проверок на чётность, выполненных в процессе вычисления LLRи декодирования, называют внешним сообщением.

### 2.2 Алгоритм «min-sum»

Алгоритм «min-sum» это модификация алгоритма «Sum-Product», получается путем упрощения вычисления сообщений от проверочных узлов к битовым. Для этого применено упрощение расчета LLRсуммы по модулю 2 множества случайных независимых величин. Это упрощает формирование сообщений от проверочных узлов к битовым. То есть вычисление их значения, теперь сводится к применению простых операций поиска минимума и сложения.

Выигрывая в простоте декодирования алгоритм серьезно проигрывает в энергетическом плане, порядка 0,2...0,5 дБ, в сравнении с «Sum-Product». Из-за чего задача в повышении эффективности алгоритма остается актуальной.

Алгоритм «min-sumnormalized» состоит в том, что теперь к расчету сообщений от проверочных узлов к битовым добавляется один элемент, так называемый коэффициент нормализации  $\alpha$ , чье значение варьируется в пределах (0,1]

Пример не единичен и для повышения эффективности алгоритма «min-sum» применяют разные модификации, в том числе и комбинации данных улучшений. Об эффективности декодеров можно судить по результатам их моделирования на большом числе пакетов и различных значения уровня шума. Моделирование проводил в своих исследованиях А.А. Хлынов [7], на рис. 4, показаны их результаты: вероятность битовой и пакетной ошибки. Так же для сравнения приведена эффективность JPL-декодера от авторов стандарта [8] (точные характеристики алгоритма и число итераций неизвестно).

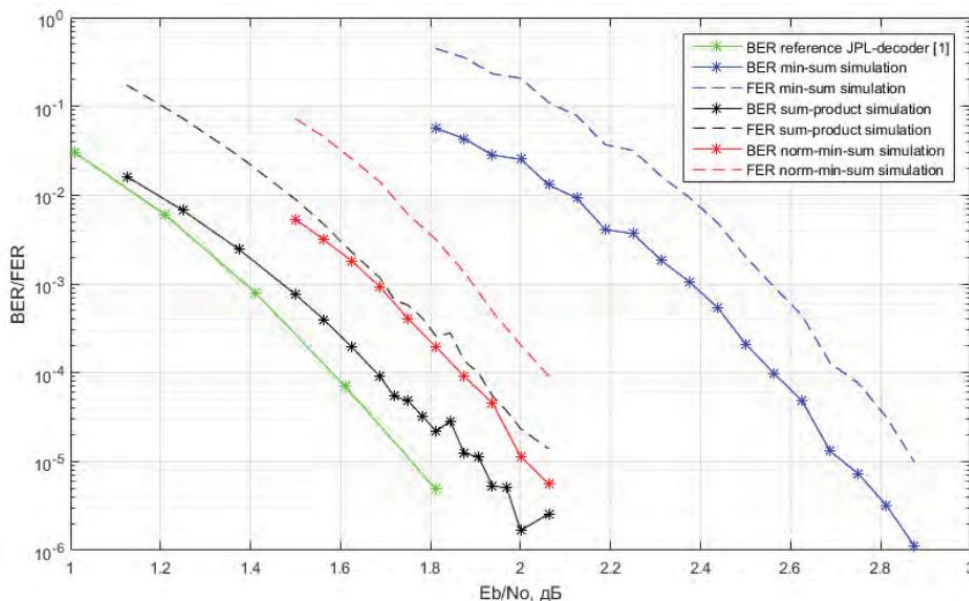


Рис.4. Вероятность битовой (BER) и кадровой (FER) ошибки для декодеров.



## Заключение

Алгоритм последовательного исключения (SC)– позволяет добиться минимальной вероятности ошибки при декодировании классических полярных кодов методом последовательного исключения. При его использовании вероятность ошибки заметно превышает вероятность ошибки декодирования по методу максимального правдоподобия. Главным недостатком алгоритма является невозможность исправления ошибок, допущенных на ранних этапах его работы.

Списочный алгоритмом последовательного исключения – это модификация алгоритма Тала–Варди. Он дает возможность использовать списочный декодер в каскадных конструкциях и позволяет осуществить эффективную аппаратную реализацию на платформах с целочисленной арифметикой.

Стековый алгоритмом последовательного исключения – его вычислительная сложность заметно ниже, чем у списочного.

Декодирование методом направленного поиска–алгоритм, чья корректирующая способность оказывается такой же, как у списочного алгоритма последовательного исключения с аналогичной размерностью списка, в котором не используется эвристическая функция.

Последовательное декодирование полярных кодов с ядром Арикана – алгоритм, чья корректирующая способность совпадает с таковой у списочного алгоритма Тала-Варди. Однако следует обратить внимание, что корректирующая способность этого кода ограничена малым минимальным расстоянием, и увеличение размера списка  $L$  не приводит к ее улучшению. При этом последовательный алгоритм декодирования полярных кодов с ядром Арикана обладает существенно меньшей сложностью по сравнению с существующими списочными и стековыми алгоритмами декодирования полярных кодов. Снижение сложности достигается за счет незначительного ухудшения корректирующей способности.

Алгоритм «Sum-Product»– алгоритм, который минимизирует вероятность ошибки на бит. Имеет высокую вычислительную сложность: на каждой итерации происходит вычисление сообщений, происходящее с целыми числовыми, а так же вычисление не тривиального вида функции.

Алгоритм «min-sum»– упрощение алгоритма «Sum-Product» в следствии чего имеет меньшую вычислительную сложность по сравнению со своим прообразом, однако остается по-прежнему сложнее алгоритмов мажоритарного декодирования и алгоритма с введением стираний из-за необходимости производить арифметические операции.

Алгоритм «min-sumnormalized» – более точен, чем алгоритм «min-sum», но по корректирующей способности уступает классическому алгоритму, хотя по легкости декодирования и выигрывает у него.

Проведенный анализ показывает текущую эффективность полярных и LDPCкодов. В ходе сравнения помехоустойчивых кодов установлено, что ни одного из приведённых в работе методов декодирования информации не позволяет полярным кодам превзойти LDPCв корректирующей способности при работе с большими блоками данных.

Полученные в ходе анализа выводы дают возможность сформулировать оптимальный метод декодирования данных с использованием полярных или LDPCкодов, что в свою очередь позволит добиться высокой помехозащищенности информации при наличии достаточного для применения описанных корректирующих алгоритмов вычислительных мощностей.

## Библиографический список

1. **Семенов, П.К.** Практические аспекты алгоритма Тала-Варди списочного декодирования полярных кодов. - Санкт-Петербург: Информатика, телекоммуникации и управление, eISSN: 2687-0517, 2012. - 40-44 с.
2. **Arikan, E.** Channel polarization: A method for constructing capacity-achieving codes for symmetric binary-input memoryless channels // IEEE Transactions on Information Theory. 2009. Vol. 55, No. 7. P. 3051–3073
3. **Миславская, В.Д.** Методы построения и декодирования полярных кодов. - Санкт-Петербург, 2014 – 20 -34 с.

4. **Тимофеев, Г.С.** «Применение полярных кодов в современных системах связи» - Красноярск: Решетневские чтения, eISSN:1990-7702, 2016. – 297-299 с.
5. Полярные коды: от теории к практике [Электронный ресурс]URL: <http://dcn.icc.spbstu.ru/index.php?id = polar>
6. IEEE Std 802.16–2012 IEEE Standard for Air Interface for Broadband Wireless Access Systems. New York, 2012
7. **Хлынов, А.А.** «Оптимизация min-sum алгоритма декодирования» - Москва: Труды Московского физико-технического института, eISSN: 2072-6759, 2016. – 13- 17 с.
8. The Consultative Committee for Space Data Systems TM synchronization and channel coding — summary of concept and rationale // CCSDS 130.1-G-2. 2012



**МОДЕЛИРОВАНИЕ КОРРЕЛЯЦИОННЫХ ХАРАКТЕРИСТИК  
ПСЕВДОСЛУЧАЙНЫХ ПОСЛЕДОВАТЕЛЬНОСТЕЙ СГЕНЕРИРОВАННЫХ  
НЕЙРОСЕТЬЮ ДЛЯ РАЗЛИЧЕНИЯ СИГНАЛОВ**

**Р.Р. Мещеров**

ORCID: 0000-0002-6364-0481 e-mail: rasil.mesherov@gmail.com

**А.В. Семашко**

ORCID: 0000-0003-4687-9666 e-mail: semashko@nntu.ru

**Н.А. Метельков**

ORCID: 0000-0003-2189-0340 e-mail: met-nikita@mail.ru

**А.П. Зайцев**

ORCID: 0000-0003-1920-6255 e-mail: cool.sasha1997@yandex.ru

Нижегородский государственный технический университет им. Р.Е. Алексева

В докладе предлагается способ генерации псевдослучайной последовательности на основе результата работы нейронной сети. Создание последовательностей предполагает учет сигналов фазо-манипулируемых в окружающей среде и подстройкой наиболее эффективных последовательностей в отношении распознавания на фоне шумов и различных сигналов фазо-манипулируемых другими последовательностями. При помощи библиотек Tensorflow и Keras создается нейросеть с обратным распространением ошибки, для тренировки используются уже существующие псевдослучайные последовательности такие как M-последовательности коды Голда и Касами. Целью работы является создание уникальных последовательностей для улучшения различения сигналов при приеме.

**Ключевые слова:** псевдослучайные последовательности, нейросети, Tensorflow, Keras, различение сигналов.

**MODELING OF CORRELATION CHARACTERISTICS OF PSEUDORANDOM  
SEQUENCES GENERATED BY A NEURAL NETWORK FOR DIFFERENTIATION  
OF SIGNALS**

**R.R. Meshеров A.V. Semashko N.A. Metelkov, A.P. Zaytsev**

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

**Purpose:** The author provides a method for generating pseudo-random sequences for distinguishing signals using neural networks.

**Design / methodology / approach:** The neural network is trained on a dataset that includes generated M-sequences, Gold and Kassami codes using the backpropagation method using the Tensorflow and Keras machine learning libraries.

**Findings:** The neural network generates a pseudo-random sequence that has a narrow autocorrelation function and a large ratio of the main lobe to the side lobe and a small cross-correlation function with extraneous noise and signals

**Research limitations/implications:** The research is limited to a data set on which the model must be trained to create pseudo-random sequences with a small cross-relationship function.

**Originality/value:** The results can be used to further improve the signal discrimination performance.

**Key words:** pseudo-random sequences, neural networks, Tensorflow, Keras, signal discrimination.

## Теоретический анализ

В современном мире все чаще используют машинное обучение и искусственные нейронные сети для самых различных задач от прогнозирования на первый взгляд случайных событий, до восстановления и улучшения фото и видео материалов.

Нейронные сети учатся путем обработки примеров, каждый из которых содержит известные «входные данные» и «результат», формируя между ними взвешенные по вероятности ассоциации, которые хранятся в структуре данных самой сети. Обучение нейронной сети на заданном примере обычно проводится путем определения разницы между обработанным выходом сети и целевым выходом. Эта разница и есть ошибка. Обучение нейронной сети на заданном примере обычно проводится путем определения разницы между обработанным выходом сети и целевым выходом [1]. Данные особенности можно применить в радиосвязи.

Для передачи информации по каналу связи используют различные типы модуляции сигналов, одним из типов является фазоманипулирование. При помощи специальных псевдослучайных последовательностей сигнал получает возможность быть принятым на фоне окружающего шума и чужих сигналов [2].

Последовательность Задорфа-Чу, имеющие другое название последовательностью Чу — это математическая последовательность, состоящая из комплексных чисел, которая при применении к сигналу порождает новый сигнал постоянной амплитуды. При наложении на сигнал циклически смещенных версий последовательности результирующий набор сигналов, обнаруженных на приемнике, не коррелирует друг с другом [3,4].

На основе уже существующих псевдослучайных последовательностей возможно при помощи машинного обучения и нейросетей создать последовательности с аналогичными свойствами для большего их разнообразия и дальнейшего улучшения.

Любые сигналы передаются в различных средах таких как вакуум, воздушное пространство, провода и т.д. Для оценки способностей сигналов быть распознанными на фоне шумов используют модель гауссовского канала связи [5].

Гауссовский канал, или канал с аддитивным белым гауссовским шумом, это канал, в котором есть только один прямой луч между передатчиком и приемником, т.е. многолучевость полностью отсутствует. Межсимвольная интерференция в этом канале так же отсутствует, а основным источником искажений, в данном случае, является тепловой шум, генерируемый в приемнике. Этот шум, как правило, имеет постоянную спектральную плотность во всей используемой полосе частот и гауссовскую функцию плотности вероятности с нулевым средним [6].

Сигналы передаются в различных средах таких как вакуум, воздушное пространство, провода и т.д. Для оценки способностей сигналов быть распознанными на фоне шумов используют модель гауссовского канала связи [7].

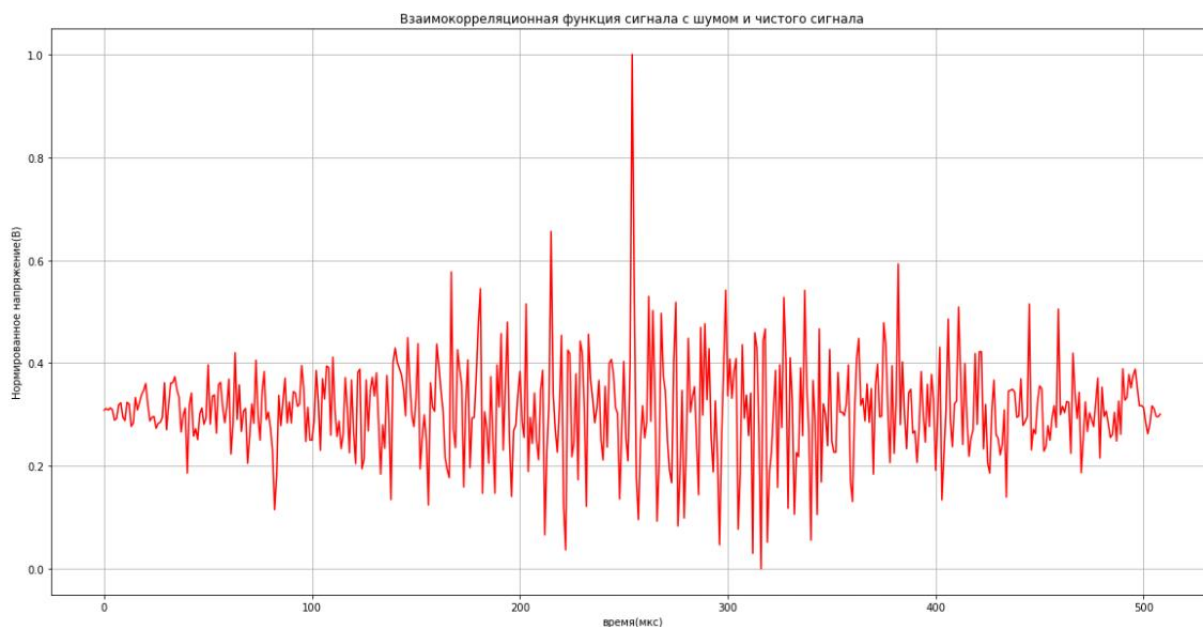
Данный тип канала часто используется в качестве канальной модели, в которой единственным нарушением связи является линейное сложение широкополосного или белого шума с постоянной спектральной плотностью и гауссовским распределением амплитуды [8]. Модель не учитывает затухание, частотную селективность, интерференцию, нелинейность или дисперсию. Тем не менее, он создает простые и поддающиеся обработке математические модели, которые полезны для понимания основного поведения системы до рассмотрения этих других явлений.

Подразумевается, что мощность сигнала падает с ростом расстояния также, как при распространении в свободном пространстве. В модели свободного пространства область между антеннами передатчика и приемника предполагается свободной от объектов, которые могли бы поглощать или отражать энергию на радиочастотах. Атмосфера, внутри этой области, ведет себя как однородная непоглощающая среда. Принимается, что земля находится бесконечно далеко от распространяемого сигнала (пренебрегаем коэффициентом отражения) [8].

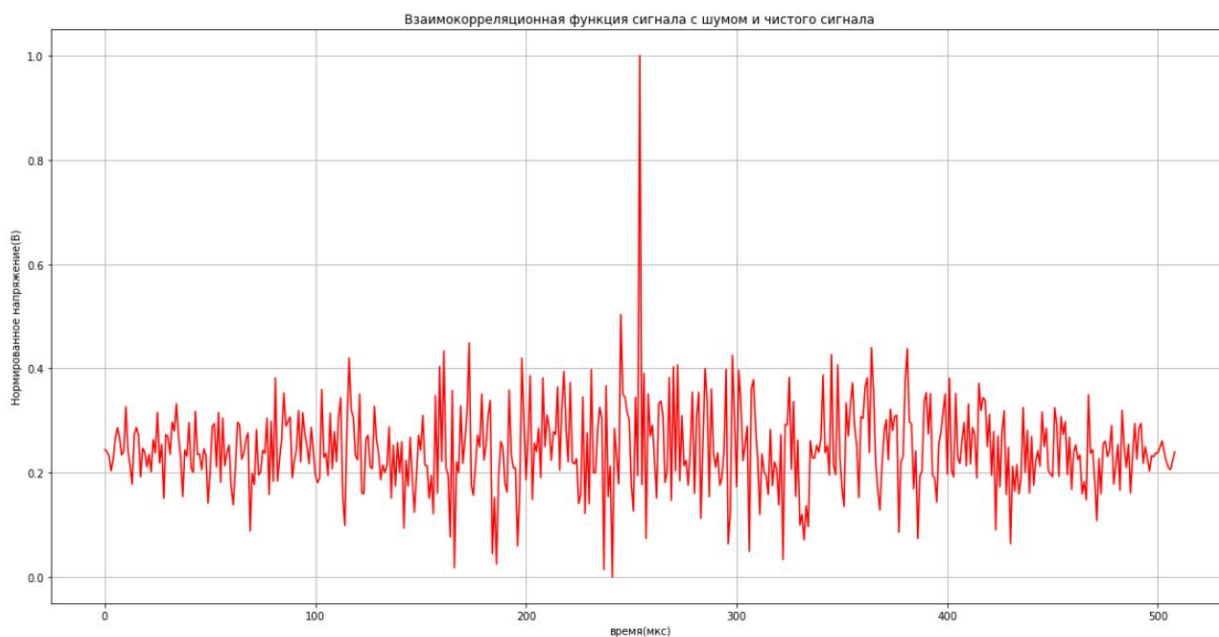
## Методика

Приведенный ниже метод будет выполнен, чтобы получать значения автокорреляционных функций последовательности. Для реализации были выбраны библиотеки на языке Python. Вычисление значений было выполнено при помощи NumPy и SciPy для генерации значений последовательности Задорфа-Чу. Для генерации псевдослучайной последовательности при помощи нейросети сначала были сгенерированы данные для обучения, являющиеся большим количеством M-последовательностей, кодов Голда и кодов Касами также при помощи вышеперечисленных библиотек. При помощи библиотек Keras и TensorFlow производится обучение модели на многослойном перцептроне с обратным распространением ошибки. За отображение результатов были ответственны библиотека Matplotlib. Данный набор позволяет работать эффективно с большими массивами значений и выводить результат в наглядной форме при помощи графиков, которые легко настраиваются.

Рассмотрим графики взаимокорреляционной функции входящего сигнала с последовательностью Задорфа-Чу и сгенерированной нейросетью секвенцией на рисунках 1 и 2 соответственно. Несущий сигнал был усилен до амплитуды 15 В и дальностью до источника 1 м. Как видно из графика автокорреляционная функция имеет явный пик, что свидетельствует о детектировании сигнала. Амплитуда шума была 60 В.

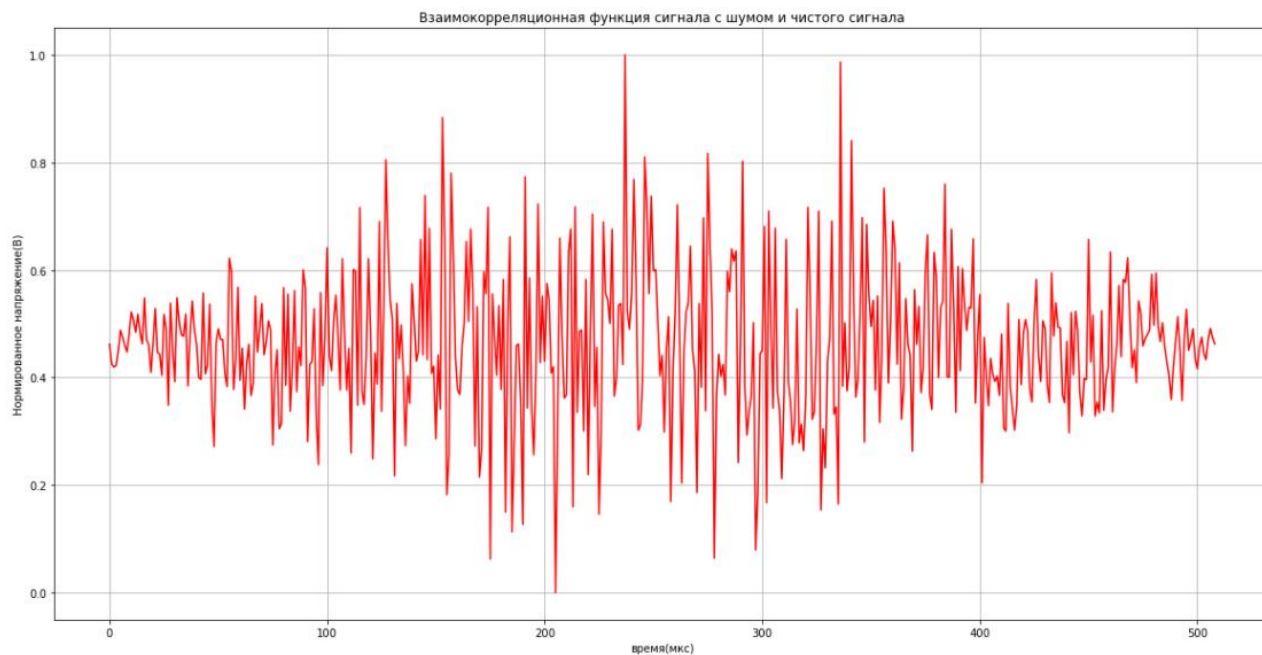


**Рис. 1. График взаимокорреляционной функции сигнала с шумом последовательности Задорфа-Чу при ее длине равной 255 с усилением в 15 раз.**

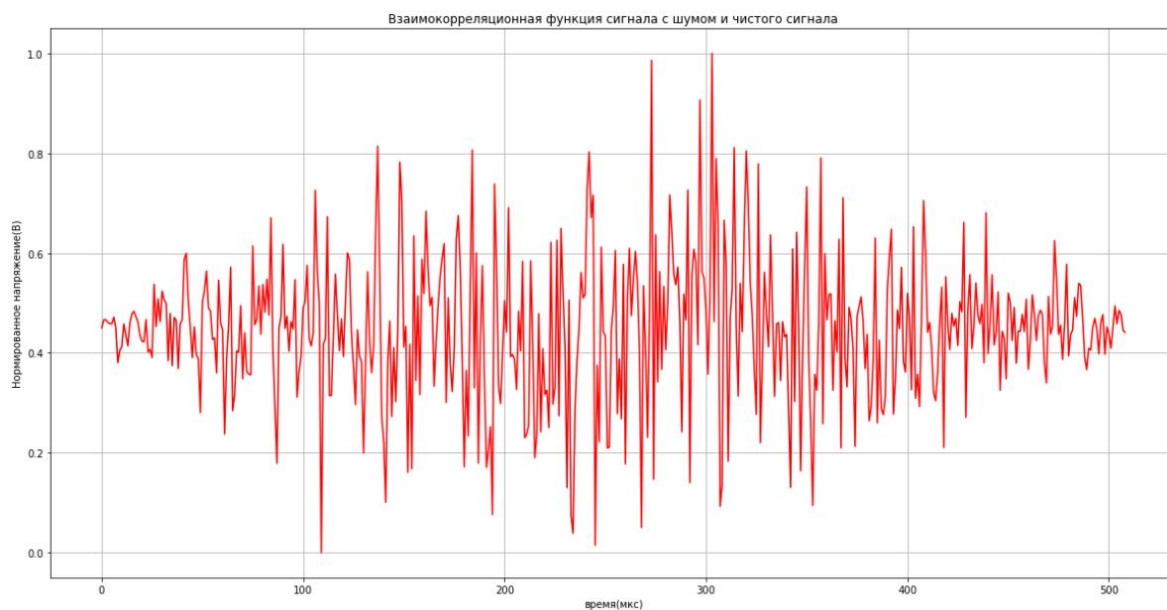


**Рис. 2. График автокорреляционной функции последовательности сгенерированной нейросетью при ее длине равной 255 с усилением в 15 раз.**

Далее на рисунках 3 и 4 представлены также графики взаимокорреляционной функции входящего сигнала с последовательностями Задорфа-Чу и сгенерированной нейросетью. Несущий сигнал остался с усилением амплитуды до 15В., но дальностью до источника, возросшей до 3 м. Как видно из графиков взаимокорреляционной функции не имеют явный пик, что свидетельствует о не обнаружении сигнала.

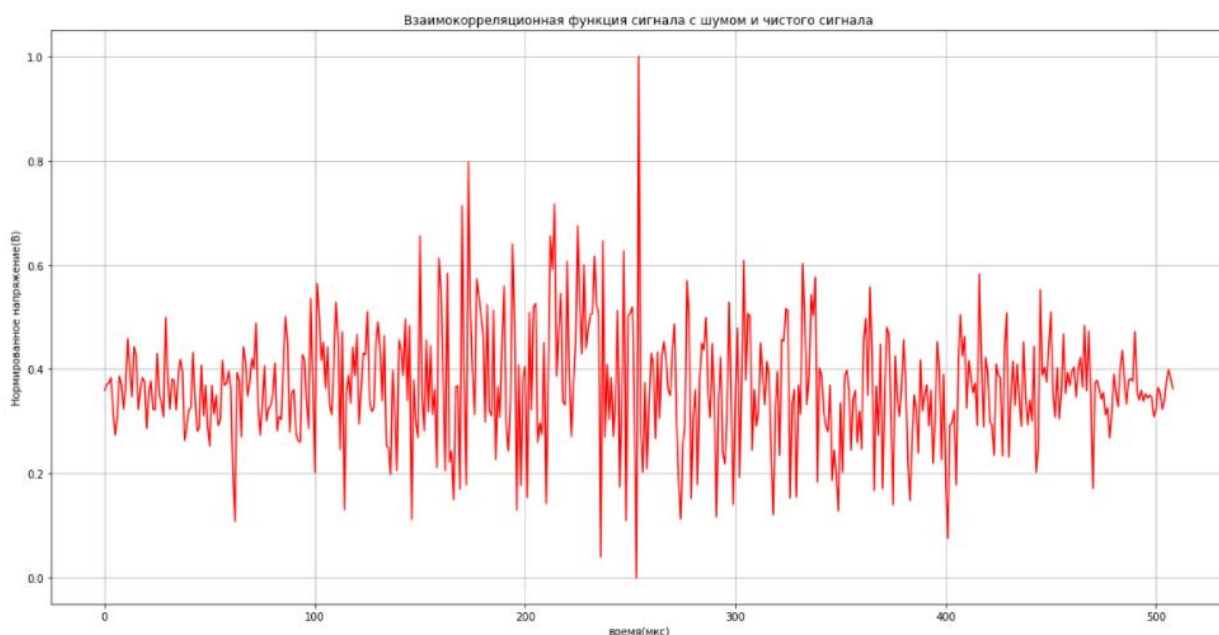


**Рис. 3. График взаимокорреляционной функции сигнала с шумом последовательности Задорфа-Чу при ее длине равной 255 с увеличенным расстоянием в 3 метра и усилением амплитуды до 15 В.**



**Рис. 4. График автокорреляционной функции последовательности сгенерированной нейросетью при ее длине равной 255 с увеличенным расстоянием в 3 метра и усилением амплитуды до 15 В.**

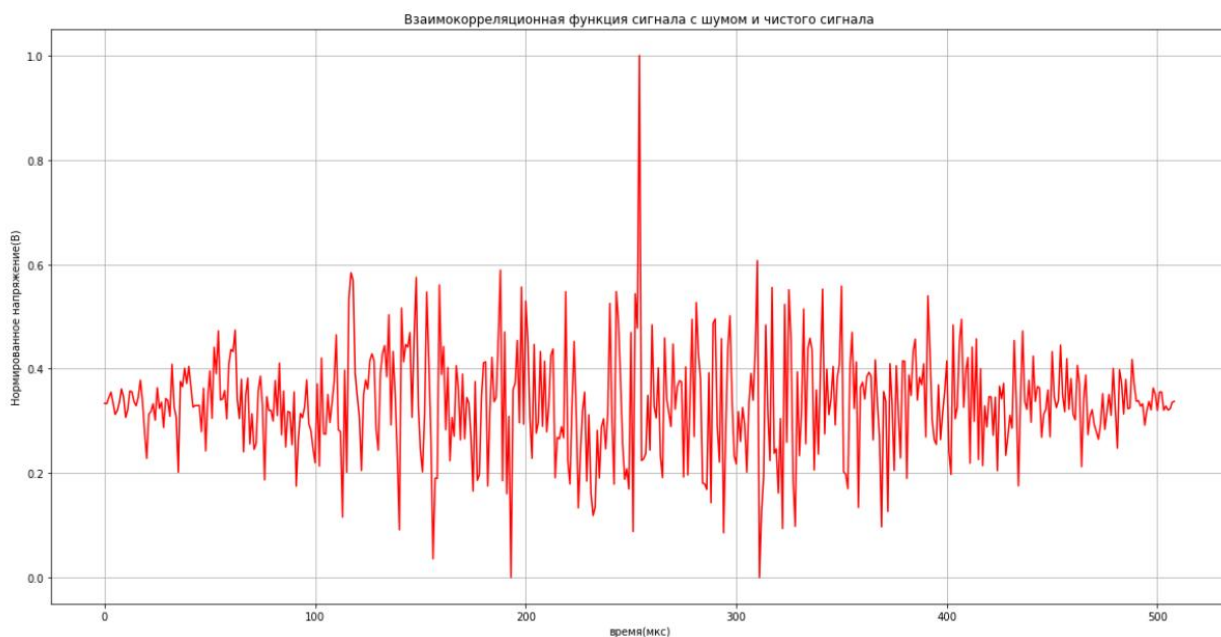
На последних рис. 5 и 6 изображены график взаимокорреляционной функции входящего сигнала и последовательностью Задорфа-Чу. Несущий сигнал был в очередной раз усилен до амплитуды 100 В и дальность до источника 3 м. На графиках вновь присутствует основной пик, который свидетельствует о обнаружении сигнала. Необходимо отметить что данные результаты получаются примерно в 75 процентов случаев используя последовательность Задорфа-Чу и примерно 85 процентов случаев используя сгенерированную нейросетью секвенцию.



**Рис. 5. График автокорреляционной функции последовательности сгенерированной нейросетью при ее длине равной 255 с увеличенным расстоянием в 3 метра и усилением амплитуды до 100 В.**

Из вышеперечисленных результатов можно пронаблюдать, зависимости между усилением несущего сигнала и его приемом, а также расстоянием от источника. Для обнаружения сигналов необходимо усиливать амплитуду их несущей. В

вышеперечисленных примерах для устойчивого обнаружения сигналов на фоне белого шума с амплитудой 60 В.



**Рис. 6. График автокорреляционной функции последовательности сгенерированной нейросетью при ее длине равной 255 с расстоянием 3 метра и усилением амплитуды до 100 В.**

Созданная нейросеть генерирует сравнительные результаты по обнаружению сигнала на фоне белого шума. Использование нейросетей позволяет увеличить количество секвенций что имеют узкую автокорреляционную функцию, и хорошо подходят для детектировании сигналов.

В результате можно сказать, что использование технологий машинного обучения в передаче информации имеет важность в улучшении качества и надежности радиосвязи.

#### **Библиографический список**

1. **Орельен, Жерон.** Прикладное машинное обучение с помощью Scikit-Learn и Tensor Flow. Концепции, инструменты и техники для создания интеллектуальных систем, 2-е изд., 2020. – 1040 с.
2. Радиосвязь / Под ред. **О.В. Головина.** - М.: Горячая линия-Телеком, 2001. - 288 с., ил.
3. **Chu, D. C.** Polyphase codes with good periodic correlation properties. IEEE Trans. Inf. Theory. 18 (4):1972. P.531–532.
4. **Beume, S. and Leung, C..** Efficient computation of DFT of Zadoff-Chu sequences. Electron. Lett. 45 (9): 2009.P.461–463.
5. **Vijay, K. Garg.** Frequency selective radio channels using power measurements/ Wireless Communications & Networking,31(4) , 2007.
6. **Witrisal, K., Kim, Y.-H., & Prasad, R.** (2001). A new method to measure parameters of frequency selective radio channels using power measurements. IEEE Transactions on Communicaitons, 49(10), 1788–1800.
7. **Скляр, Б.** Цифровая связь. Теоретические основы и практическое применение/ Digital Communications: Fundamentals and Applications. - 2-е изд. -М.: Вильямс, 2007. - 1104 с.
8. **Ферр, К.** Беспроводная цифровая связь. Методы модуляции и расширения спектра/ Wireless Digital Communications: Modulation and Spread Spectrum Applications. – М.: Радио и связь, 2000. – 552 с.

## КОНЦЕПЦИЯ И РЕАЛИЗАЦИЯ СИСТЕМЫ ПОСТРОЕНИЯ ЦИФРОВЫХ КАРТ МЕСТНОСТИ, НА ОСНОВЕ ЭКСПЕРТНОЙ СИСТЕМЫ ВЫБОРА ОПТИМАЛЬНОЙ КАРТОГРАФИЧЕСКОЙ ПРОЕКЦИИ

К.О. Староверова

ORCID:0000-0001-6463-676X e-mail: kristi-star-18@yandex.ru

Нижегородский государственный технический университет им. Р.Е. Алексеева

Цель доклада заключается в рассмотрении особенностей разработки информационной системы построения цифровых карт местности, используемых при работе радиолокационной станции. Значительное внимание уделяется детальному анализу и синтезу методики выбора оптимальной по заданным критериям картографической проекции. На базис синтезированной методики, представлены этапы реализации экспертной системы, включая выбор средств её разработки и составления базы знаний на основе научно установленных фактов и информации. Разработанная экспертная система позволяет значительно ускорить процесс выбора той картографической проекции, которая по всем параметрам наиболее точно удовлетворяет условиям функционирования радиолокационной станции и обеспечивает минимальные неточности. В разработанную систему построения цифровых карт заложен математический аппарат картографических проекций, а также механизм представления текстовой информации о карте в виде изображения. В статье представлены результаты проверки корректности взаимодействия экспертной системы (ЭС) и системы построения цифровых карт местности (ЦКМ), а также правильности представления карты путём сравнения с её изображением в существующих геоинформационных системах. Совместная работа ЭС и системы построения ЦКМ позволяет наиболее грамотно и оперативно проектировать карты для отображения на них радиолокационной информации.

**Ключевые слова:** картографическая проекция, цифровая карта местности, экспертная система, радиолокационная станция.

## THE CONCEPT OF IMPLEMENTING A SYSTEM FOR CONSTRUCTING DIGITAL MAPS OF THE AREA, BASED ON THE EXPERT SYSTEM FOR CHOOSING THE OPTIMAL CARTOGRAPHIC PROJECTION

K.O. Staroverova

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

**Purpose:** The purpose of the scientific article is to consider the features of the development of an information system for constructing digital terrain maps used in the operation of a radar station.

**Design/methodology/approach:** The article examines the methodology for determining the optimal cartographic projection. Based on the analyzed methodology, the stages of developing an expert system for choosing the optimal cartographic projection are determined, which include the choice of development tools and the creation of a knowledge base and rules. Much attention is paid to the study of the operation of the digital terrain mapping system and its interaction with the expert system.

**Conclusions:** The expert system allows you to speed up the process of choosing a cartographic projection and determines the one that most accurately satisfies the conditions of the radar station in all parameters. The system of building digital terrain maps together with the expert system allows you to design maps most competently and quickly to display radar information on them.

**Consequences of research/limitations:** In the future, the study can be supplemented by adding new cartographic projections, expanding the knowledge base, using a new information display mechanism.

**Originality/value:** The system can be applied at enterprises of the military-industrial complex. It has been studied that there are few developments of a similar nature in the digital market and they are inaccessible.

**Keywords:** cartographic projection, digital terrain map, expert system, radar station.



В настоящее время методы и средства радиолокации получили широкое применение во многих областях гражданских и военных разработок. Однако, основное их назначение заключается в исследовании территории страны и её воздушного пространства на предмет безопасности. Для этой цели используются радиолокационные станции (РЛС), которые устанавливаются на различных передвижных устройствах, кораблях и летательных аппаратах. Радиолокационные станции позволяют обнаруживать, идентифицировать и распознавать различные цели в любое время суток, практически при любой погоде и на весьма значительных расстояниях. Одной из основных задач РЛС является сопровождение целей, то есть определение с высокой точностью координат целей, параметров их движения, таких как значения дальности до них, их угловые координаты, также вектора скоростей и ускорений.

Указанная задача реализуется благодаря встроенным в РЛС устройствам, таким как передающая и приёмная антенны, радиоприёмник, передатчик и выходное интеллектуальное устройство, которое передает результат их совместной деятельности на индикатор кругового обзора. Данное выходное устройство предназначено для отображения радиолокационной информации на экране индикатора в координатах азимут-дальность на фоне радиально-круговой развёртки. Под ней отображается цифровая карта местности (ЦКМ), которая содержит информацию о территории, необходимую для работы станции. Все данные, представленные в индикаторе кругового обзора, в совокупности дают оператору РЛС наглядное представление об окружающей обстановке.

Конечной целью всех вышеперечисленных задач является максимально точное определение положения выявленных целей на земной поверхности. Положение этих объектов может быть определено в различных системах координат[1]. Однако, все процедуры для выполнения расчётных заданий ведутся в геодезической и плоской прямоугольной системе координат. Прямоугольная система является более удобной для практического использования. Но её применение вместо геодезических координат требует перехода от поверхности земного эллипсоида на плоскость, что выполняется за счёт использования картографических проекций[6]. Ряд картографических проекций широк и при этом каждая из них подходит для разных территорий, а также вносит различные величины искажений в отображение информации, что формирует ошибочное представление.

Одним из отличительных свойств радиолокационной станции является её мобильность. Это означает, что при изменении территориального положения РЛС, необходимо проверить, подходит ли созданная карта для текущего местонахождения. В условиях работы РЛС, когда необходимо быстро принимать решения, вручную это делать нецелесообразно (трудоемко, времязатратно, не точно). Поэтому существует потребность в разработке экспертной системы (ЭС), автоматически подбирающей картографическую проекцию для отображения цифровой карты местности (ЦКМ) в индикаторе кругового обзора РЛС.

Для обеспечения точности и корректности работы такой системы, она должна анализировать различные факторы, влияющие на выбор картографической проекции, к числу которых относятся: факторы, характеризующие объект картографирования (территория); факторы, характеризующие получаемую проекцию (искажения)[5]. Однако, обработка критериев, характеризующих отображаемую территорию, является первостепенной, так как определение характера искажений, вносимых проекцией, основывается исключительно на математических расчетах на базе научно установленных формул. В связи с этим, критерии выбора картографической проекции по описываемой территории имеют особо важное значение и рассмотрены в данной статье. Для выполнения этой задачи, эффективно применить экспертную систему, которая путем работы с заложенной в ней базой знаний, заменит специалиста-эксперта и выдаст пользователю (оператору РЛС) некоторую, объективно приемлемую группу проекций. Из данной группы проекций, в дальнейшем, необходимо выбирать оптимальную проекцию по критерию минимума картографических



искажений.

Экспертная система выбора группы проекций, должна оперировать такими параметрами как:

1. **Географическое положение территории.** Для изображения территорий, расположенных в приэкваториальной зоне (приблизительно симметрично относительно экватора), выбирают цилиндрические проекции; для полярных зон – азимутальные; для территорий, расположенных в средних широтах – конические[6].
2. **Геометрическая форма территории.** Для изображения территорий, имеющих примерно округлую форму, применяют азимутальные проекции в различных аспектах(нормальные, косые, поперечные). Для территорий, вытянутых вдоль меридианов или под углом к ним, используют поперечные или косые цилиндрические проекции[2].
3. **Территориальный охват территории.** По территориальному охвату выделяются картографические проекции для карт мира, полушарий, материков и океанов, карт отдельных государств и их регионов[1].

Таким образом, при выборе проекции в ходе работы с экспертной системой, необходимо решить следующие вопросы:

1. Определить местоположение области отображения относительно экватора (полярная, среднеширотная, экваториальная). Местоположение определяется точкой стояния РЛС: если ее широта находится в диапазоне от  $0^{\circ}$  до  $23^{\circ}26'22''$  или от  $0^{\circ}$  до  $-23^{\circ}26'22''$ , то область отображения является экваториальной; если от  $23^{\circ}26'22''$  до  $66^{\circ}33'39''$  или от  $-23^{\circ}26'22''$  до  $-66^{\circ}33'39''$ , то это среднеширотная область; если от  $66^{\circ}33'39''$  до  $90^{\circ}$  или от  $-66^{\circ}33'39''$  до  $-90^{\circ}$ , то это полярная область. При этом, если точка расположена на границе областей, что составляет  $\pm 5^{\circ}$  от указанных значений, целесообразно учитывать проекции, предназначенные для обеих соседних областей.
2. Определить форму или направление области отображения (квадратный участок, направлен с востока на запад или с севера на юг). Форма отображаемой территории определяется на основе минимальных и максимальных значений широты и долготы карты. Если диапазон широт равен диапазону долгот, с разницей до  $\pm 5^{\circ}$ , то форма территории считается квадратом. Если диапазон широт превышает диапазон долгот более чем на  $\pm 5^{\circ}$ , то форма территории считается вытянутой с севера на юг. Иначе, если диапазон долгот превышает диапазон широт, то территория вытянута с запада на восток.
3. Определить размер области отображения (мир, полушарие, континент, более мелкие регионы). Размер отображения территории также определяется в зависимости от минимальных и максимальных значений широты и долготы. Например, если диапазон широт от  $-90^{\circ}$  до  $90^{\circ}$  и долгот от  $-180^{\circ}$  до  $180^{\circ}$  то объем отображения – мир. Однако, если диапазон широт и долгот составляют  $50^{\circ}$  и менее, то объем отображения – регион.

Как было сказано ранее, экспертная система работает на основе базы знаний, которая содержит «правила» вывода и обработки информации, а также данные об опыте человека и его знаниях. Наиболее классическим способом представления «правил» являются продукционные правила[3]. Продукционное правило – это правило вида: *ЕСЛИ* <условие> *ТО* <действие>. Формулируя на естественном языке, в рассматриваемой предметной области можно сформулировать следующее правило: «*ЕСЛИ* область рассмотрения не большая, расположена на экваторе и имеет квадратную форму, *ТО* оптимальной проекцией будет проекция Гаусса-Крюгера».

Составные части «правила» несут в себе некие факты, состоящие из «переменной» и «значения», которое берется из области возможных значений. То есть одно из «условий» «область рассмотрения не большая» можно переписать в виде: «Объем отображения = регион», где «Объем отображения» – переменная, а «регион» – её значение. Если все возможные вопросы относительно «объема отображения» будут иметь ответы «мир», «полушарие», «континент», «регион» – то совокупность ответов будет являться доменом этой переменной.

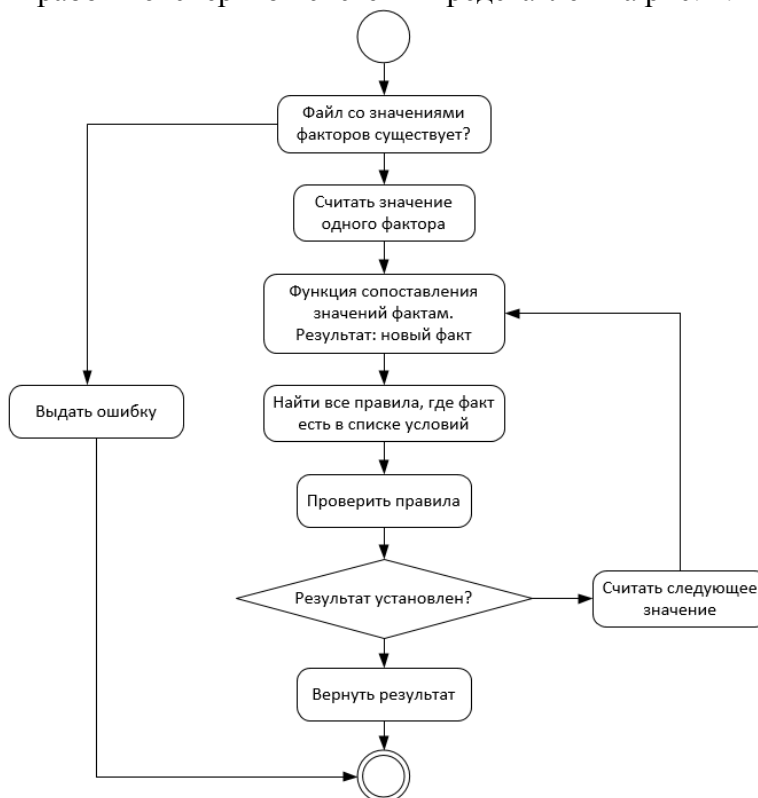
«Действие» представляет собой присвоение «истинности» некоторому «факту». То есть вместо описания на естественном языке «оптимальной проекцией будет проекция Гаусса-Крюгера», в ходе работы «правила» факт «*Оптимальная проекция = проекция Гаусса-Крюгера*» помечается как «истинный». Таким образом, из рассматриваемого примера получается правило: ЕСЛИ «объем отображения = регион» И «расположение = средние широты» И «ориентация = север/юг» ТО «оптимальная проекция = проекция Гаусса-Крюгера».

Для реализации экспертной системы используется среда CLIPS, работающая на основе продукционных правил и целеориентированная на создание экспертных систем, благодаря чему позволяет грамотно записать правила вывода. На вход ЭС поступает файл (рис. 1), с заданными значениями факторов, указанных при рассмотрении вопросов выбора картографической проекции. Данные значения определяются программно, путем обработки файлов, хранящих информацию об отображаемой территории.

Файл	Правка	Формат	Вид	Справка
1:1000000				//масштаб карты
59.99993; 43.9972128				// территориальное расположение РЛС
72.0000108; 23.9960865				
72.0000108; 59.9979348				
43.9998278; 35.997021				//геометрическая форма и
43.9998278; 51.9969565				//терриальный охват

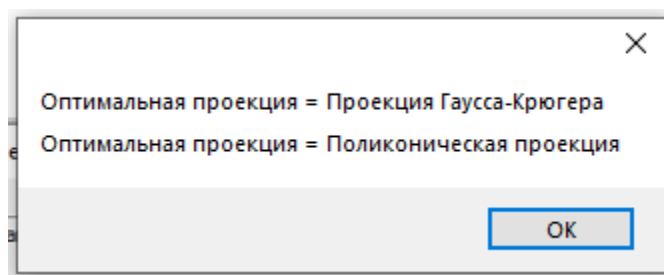
**Рис.1. Вид файла, хранящего значения факторов выбора проекции**

Механизм логического вывода считывает из файла значения факторов, проверяет и ставит в соответствие некоторый факт. После этого, на основе этих фактов, в базе знаний находятся все правила, в которых этот факт(ы) содержится в качестве условия. Если такое правило существует и все его условия являются истинными, то условие из левой части правила добавляется к сработавшим фактам и если, правило завершающее, то выводится результат. Алгоритм работы экспертной системы представлен на рис. 2.



**Рис. 2. Алгоритм работы экспертной системы**

После того, как все факты приняты в рассмотрение, а правила проверены, выводится результат работы ЭС. В рассматриваемом примере, из заданных проекций, были выбраны две, которые действительно подходят под заданные параметры (рис. 3).



**Рис.3. Результат работы экспертной системы**

Таким образом, экспертная система, подключенная к системе отображения радиолокационной информации, позволит получать названия проекций из сформированной группы, для построения цифровой карты местности.

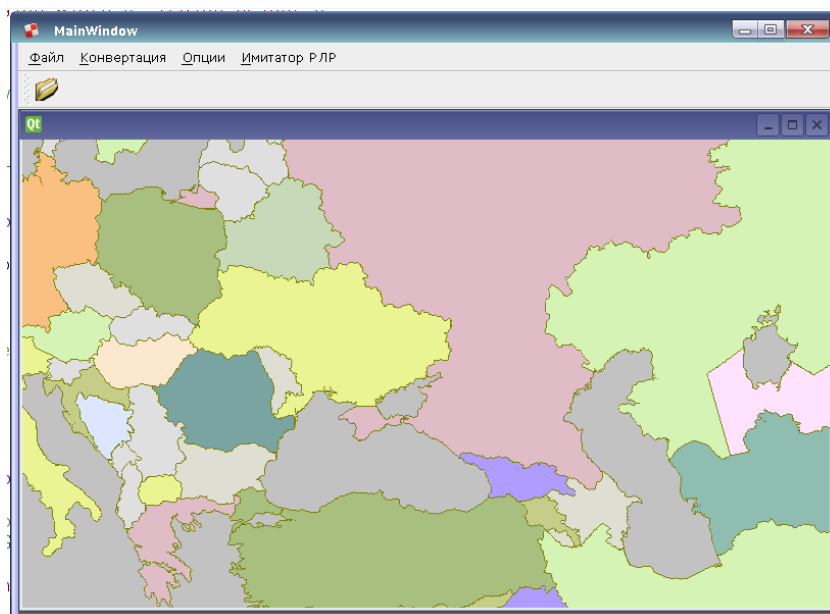
Создание карты выполняется на основе векторных данных, объектами которых являются точки, полилинии (ломаные), полигоны (многоугольники). Эти данные хранятся в файлах, представленных в формате MIF-MID, который является открытым текстовым обменным форматом. Пример файла в формате MIF-MID, содержащего информацию о точечном слое представлен ниже. Формат предполагает наличие двух файлов:

1. MIF – описание проекций, трансформаций и прочих дополнительных сведений, а, главное, перечень пространственных объектов (точки, линии, полигоны) с их координатами.
2. MID – одноименный с MIF файл, в котором для каждого пространственного объекта, описанного в MIF, приводится строка с семантическими данными.

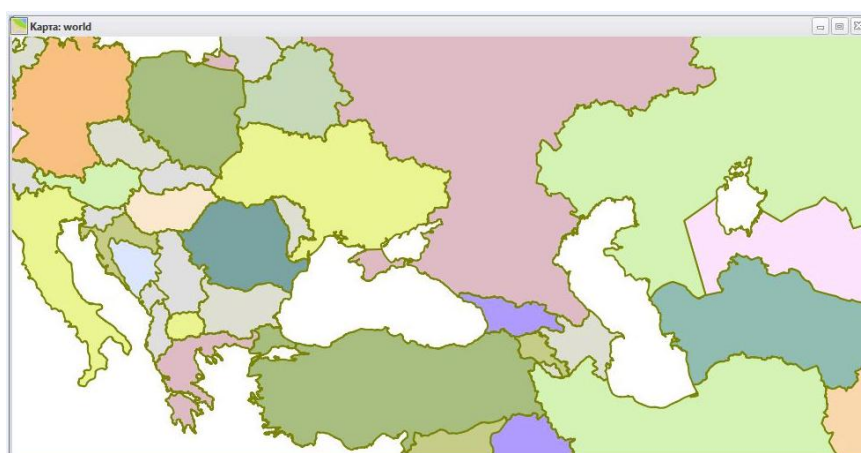
Для разработки программного обеспечения используется язык программирования C++ и фреймворк для его применения Qt. Qt позволяет запускать написанное с его помощью программное обеспечение в большинстве современных операционных систем путём простой компиляции программы для каждой системы без изменения исходного кода.

Одним из условий для работы программы, является то, что работа с картой выполняется посредством применения данных, хранящихся в файле формата mgl. При этом по умолчанию, все цифровые карты представлены форматом mif-mid. Для того, чтобы выполнить поставленное условие, необходим конвертор данных из одного формата в другой. Конвертор в качестве входных данных принимает файл формата mif, выполняет чтение и обработку данных из него и на выход предоставляет файл формата mgl. Файл mgl содержит информацию о точках карты в формате прямоугольных координат X и Y, в отличие от mif-mid в котором хранятся геодезические координаты B и L. Поэтому при конвертации данных применяется картографическая проекция, которая выбирается на основе рекомендаций экспертной системы. Её применение означает передачу значений координат B, L всех объектов в соответствующие проекции формулы.

Для того, чтобы проверить правильность составления системой карты в выбранных экспертной системой проекциях, необходимо сравнить её с картой, составленной в геоинформационной системе. На рисунках 4 и 5 отображены карты в созданной системе и в ГИС «Аксиома», соответственно, построенные в проекции Гаусса-Крюгера.



**Рис.4. Вид карты построенной в системе отображения**



**Рис.5. Вид карты, построенной в ГИС Аксиома**

Как видно из рисунков, карты похожи по расположению объектов. Однако, для их сравнения необходимо использовать более точные методы, одним из которых является метод сопоставления плоских прямоугольных координат произвольно выбранных точек на обеих картах. Результаты сравнения представлены в табл. 1, где в 1 столбце расположены координаты точек, в значениях широты и долготы на поверхности Земли, а во 2 и 3 координаты этих точек, переведенные в прямоугольные координаты, разработанной системой и ГИС Аксиома.

**Таблица 1.**

**Сравнения прямоугольных координат в созданной системеи ГИС «Аксиома»**

	<b>Разработанная система</b>	<b>ГИС «Аксиома»</b>
<b>(В, L) – в градусах</b>	<b>(X, Y) – в метрах</b>	<b>(X, Y) – в метрах</b>
(53.4931, 18.4877)	(6020353.953980, 5540273.909724)	(6032535.745, 4333276.878)
(52.8412, 45.5823)	(5923533.331779, 7345640.436504)	(5957200.132, 7539232.603)
(44.5888, 33.5224)	(4929704.980219, 6541467.761429)	(4939478.464, 6541491.712)
(42.9056, 25.8003)	(4767578.195586, 5912137.679735)	(4753010.191, 5402021.617)
(35.8217, 45.5148)	(4028795.696427, 7633308.081919)	4065886.206, 7546514.819)
(34.8758, 35.9834)	(3854837.177319, 6772700.180746)	(3864949.621, 6772801.38)
(33.6438, 38.7275)	(3728940.190344, 7031475.355700)	(3724261.276, 7474720.119)

Исходя из данных, приведённых в табл. 1, можно сделать вывод, что разработанная система отображения радиолокационной информации правильно рассчитывает координаты объектов и строит цифровые карты местности. Таким образом, подключенная к РЛС экспертная система позволяет значительно улучшить качество работы РЛС. Кроме того, экспертная система может быть легко усовершенствована, расширена путём увеличения её базы знаний, что повысит точность представления данных.

#### **Библиографический список**

1. **Snyder, J. P.** Map Projections: A Working Manual. U.S. Geological Survey Professional Paper 1395. Washington, DC: United States Government Printing Office, 1987.
2. **Бобров, И.Н., Воронов, Г.Б.,** Исследование путей выбора оптимальных проекций при заданных условиях отображения картографической информации в геоинформационных системах военного назначения // Геодезия и картография. – 2012. – № Спецвыпуск. – С. 150-153. DOI: 10.22389/0016-7126-2012-150-153
3. **Любарский, Ю.Я.** Интеллектуальные информационные системы / Ю.Я. Любарский. - М.: Наука, 2015. - 228 с.
4. **Частиков, А.П.** Разработка экспертных систем. Среда Clips. / А.П.Частиков, Т.А.Гаврилова, Д.Л.Белов.–Спб.: "БХВ-Петербург". – 2003 – 395с.
5. **Запороженко, А.В.** Картографические проекции и методика их выбора для создания карт различных типов. – Ногинск: Панорама, 1991-2007.
6. **Топчилов, М.А.** Картография: учебно-метод. пособие. Издание 2-е, перераб. и доп. [Текст] / М.А. Топчилов, Л.А. Ромашова, О.Н. Николаева. – Новосибирск: СГГА. – 2009. – 109 с.
7. **Утешева, Т.Ш.** Математическое проецирование в ГИС: учебно-методическое пособие / Т.Ш. Утешева. – Нижний Новгород: "Нижегородский госуниверситет", 2016. – 60 с.

## РАЗРАБОТКА СРЕДСТВ АВТОМАТИЗАЦИИ ДЛЯ СПЕЦИФИКАЦИИ И УПРАВЛЕНИЯ ТРЕБОВАНИЯМИ К ПРОГРАММНОМУ ОБЕСПЕЧЕНИЮ ИНФОРМАЦИОННЫХ СИСТЕМ

А.А.Градиленко

ORCID: 0000-0003-2410-9900 e-mail: artem.gradilenko@gmail.com

Ю.С. Егоров

ORCID: 0000-0002-6628-4574 e-mail: ckar@list.ru

Т.И. Горячева,

ORCID: 0000-0003-4668-4156 e-mail: gor43210@yandex.ru

Нижегородский государственный технический университет им. Р.Е.Алексеева

Доклад посвящен разработке веб-сервиса для автоматизации спецификации менеджмента требований к программному обеспечению (SMSR) информационных систем (ИС).

В статье рассматриваются процессные подходы к проектированию и внедрению информационных систем, также методы поддержки принятия решений для спецификации менеджмента требованиями к ПО (SMSR).

Разработаны требования к инструменту SMSR, варианты использования, архитектура и база данных веб-сервиса. Предложен подход к созданию инструмента автоматизации SMSR. Разработан оригинальный алгоритм поддержки принятия решений. Программное обеспечение протестировано и отлажено. Выбраны направления для будущих приложений.

Настоящая статья является отправной точкой для дальнейших исследований по созданию системы поддержки принятия решений для спецификации менеджмента требованиями к ПО на протяжении всего жизненного цикла ИТ-системы.

Веб-сервис содержит оригинальный алгоритм поддержки принятия решений, который позволяет пользователю создавать SMSR с анализом знаний других разработчиков. Содержимое разделов SMSR автоматически выбирается из базы данных похожих проектов и анализируется.

**Ключевые слова:** веб-сервис, спецификация менеджмента требованиями к ПО (SMSR), информационная система (ИС), поддержка принятия решений, жизненный цикл ПО.

## DEVELOPMENT OF AUTOMATION TOOLS FOR SPECIFICATION AND MANAGEMENT OF SOFTWARE REQUIREMENTS FOR INFORMATION SYSTEMS

A.A.Gradilenko, Y.S.Yegorov, T.I. Goruasheva

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

**Purpose:** The article is devoted to the development of a web service for automating the of information systems (IS).

**Design/methodology/approach:** The article discusses process approaches to the design and implementation of information systems, as well as decision support methods for the specification and management of software requirements (SMSR).

**Findings:** Requirements for the SMSR tool, use cases, architecture and database of the web service have been developed. An approach to creating an SMSR automation tool is proposed. An original decision support algorithm has been developed. The software has been tested and debugged. Directions for future applications are chosen.

**Research limitations/implications:** This article is a starting point for further research on the creation of a decision support system for specification and management of software requirements (SMSR) throughout the life cycle of an IT system.

**Originality/value:** The web service contains an original decision support algorithm that allows the user to create SMSR with analyzing of the knowledge of other developers. The content of SMSR sections is automatically selected from a database of similar projects and analyzing.

**Key words:** web service, specification and management of software requirements (SMSR), information system (IS), decision support, software life cycle.

## Введение

Разработка программного обеспечения (ПО) информационных систем (ИС) – сложный процесс, который состоит из множества различных процессов жизненного цикла ПО от моделирования бизнес-процессов, которые планируется реализовать в виде программного продукта, до процессов практического внедрения и сопровождения конечной системы [1, 2, 3]. Большинство технических заданий (ТЗ) на разработку информационных систем обусловлены тем, что Заказчику необходимо автоматизировать некоторые бизнес-процессы, при этом разработчик готов оказать услуги по автоматизации этих процессов, путём создания программного продукта. ТЗ является основанием для запуска технологического процесса разработки ПО.

Большинство проблем по разработке ПО ИС связаны с организационными моментами, в частности, это проблемы эффективного управления проектом – эффективной разработки технического задания и эффективного отслеживания выполнения этого задания (менеджмента требованиями). В начале нового тысячелетия многие (порядка 30%) крупных сложных проектов не дошли до своего финиша по причине потери управляемости проекта – плохого отслеживания выполнения технического задания и плохой разработки технического задания. При этом выяснилось, что даже лучшие методы и эффективные инструменты программирования теряют свои достоинства, если некорректно составлено техническое задание на проект, также, если изменения требований ТЗ по обоснованным причинам в процессе проектирования не были доведены до всех разработчиков. В результате плохой работы с требованиями накапливается постепенно высокий уровень хаоса в системе и проект разваливается не доходя до финиша. Чтобы избежать таких ситуаций очень важно использовать инструменты управления требованиями, которые позволяют более корректно и квалифицированно разработать ТЗ и в ходе проектирования держать под контролем в «ежевых рукавицах» все требования ТЗ.

Техническое задание (ТЗ) – исходный документ на проектирование технического объекта (изделия). Оно устанавливает основное назначение разрабатываемого объекта, его технические характеристики, показатели качества и технико-экономические требования, предписание по выполнению необходимых стадий создания документации (конструкторской, технологической, программной и т.д.) и её состав, а также специальные требования [1, 2, 3].

По статистике, ТЗ при каждой новой разработке пишется без учёта содержания технических заданий предыдущих разработок, что значительно увеличивает временные затраты, необходимые на его формирование, что в свою очередь увеличивает временные затраты на разработку ПО в целом. Инструмент динамического объектно-ориентированного управления требованиями DOORS (Dynamic Object-Oriented Requirement System) фирмы IBM Rational/Telelogic [4] позволяет разрабатывать спецификации требований и управлять требованиями в ходе разработки при реализации больших сложных проектов, когда многие сотни разработчиков соисполнителей разнесены по всему земному шару. DOORS имеет свой веб-сайт, который связывает всех разработчиков одного проекта.

Тем не менее, в настоящее время отсутствуют отечественные инструменты управления проектами типа DOORS, которые позволяют автоматизировать процесс создания ТЗ и отслеживать его выполнение. Поэтому перед российскими разработчиками стоит задача по импортозамещению иностранного ПО вследствие санкций. На это направлены госпрограммы: «Развитие промышленности и повышение ее конкурентоспособности», Программа «Поддержки цифровизации малого и среднего бизнеса» Минцифры России в рамках федерального проекта «Цифровые технологии» национальной программы «Цифровая экономика Российской Федерации».

По этой причине актуальным и целесообразным направлением исследований становится разработка отечественных средств автоматизации спецификации и управления требованиями к ПО ИС и создание соответствующего веб-сервиса.

Целью работы является повышение эффективности процесса формирования типового ТЗ на разработку ПО ИС за счёт повторного использования знаний и накопленного опыта признанных авторитетов в конкретных предметных областях ИС, также использования мирового накопленного опыта в виде открытых международных стандартов системной и программной инженерии. Для достижения поставленной цели требуется решение задач:

- исследования предметной области;
- выбора средств программной реализации веб-сервиса;
- проектирования архитектуры и алгоритмов функционирования веб-сервиса;
- проектирования схемы базы данных веб-сервиса;
- программной реализации веб-сервиса.

### **Исследование предметной области**

**Техническое задание**, как термин в области информационных технологий – это юридически значимый документ, содержащий исчерпывающую информацию, необходимую для постановки задач исполнителям на разработку, внедрение или интеграцию программного продукта, информационной системы, сайта, портала либо прочего ИТ сервиса.

**Техническое задание** является **юридическим документом, который** как приложение включается в договор между заказчиком и исполнителем на проведение проектных работ и является его основой: определяет порядок и условия работ, в том числе, цель, задачи, принципы, ожидаемые результаты и сроки выполнения. В ТЗ должны быть указаны объективные критерии, по которым можно определить, сделан тот или иной пункт работ или нет.

Все изменения, дополнения и уточнения формулировок ТЗ обязательно согласуются с заказчиком, им утверждаются и доводятся до исполнителей – разработчиков ПО.

Это необходимо и потому, что в случае обнаружения в процессе решения проектной задачи неточностей или ошибочности исходных данных возникает необходимость определения степени вины каждой из сторон-участниц разработки, распределения понесённых в связи с этим убытков.

### **Основные стандарты системной и программной инженерии (СПИ)**

Важнейшим компонентом методологического базиса системной и программной инженерии являются официальные международные стандарты, содержащие описание подходов и методов создания систем различных классов и назначения, а также задающие правила работы, применимые в сфере системной инженерии. Такие стандарты выделены в специальный домен. Среди официальных стандартов системной и программной инженерии (СПИ) главенствующее место сегодня занимают спецификации, разрабатываемые седьмым подкомитетом Объединённого технического комитета 1 ИСО и МЭК — Системная и программная инженерия (ISO/IEC JTC1/SC7 Software and systems engineering), а также некоторыми другими подкомитетами.

В составе стандартов СПИ традиционно выделяют четыре группы спецификаций:

- стандарты основ СПИ;
- стандарты создания систем;
- стандарты управления и оценки свойств систем и процессов;
- стандарты описания систем и процессов.

Важной характеристикой системы стандартов СПИ является целостность.

В плане повышения целостности стандартов СПИ специальный комитет №7 (SC7) совместной технической комиссии №1 (JTC1) (JTC1/SC7) проводит работы по горизонтальной, вертикальной и межотраслевой гармонизации стандартов СПИ. Горизонтальная гармонизация ведётся в разрезе процессы СПИ — зрелость процессов — качество процессов, вертикальная гармонизация ведётся в разрезе: системы в целом — домены — отдельные процессы, а межотраслевая гармонизация идёт путём согласования



планов и содержания работ по стандартизации, проводимых различными группами. Одним из важных итогов этой работы считается появление в 2008 г. гармонизированных между собой стандартов процессного типа:

- «ISO/IEC 15288:2008 — СПИ. Процессы жизненного цикла систем». Устанавливает общие принципы описания ЖЦ систем, созданных людьми. Определяет набор процессов ЖЦ и соответствующую терминологию.
- «ISO/IEC 12207:2008 — СПИ. Процессы жизненного цикла программных средств». Устанавливает, используя чётко определённую терминологию, общую систему процессов жизненного цикла программных систем (ЖЦ ПС), на которую можно ориентироваться в программной индустрии.

На сегодняшний момент это основные, базовые стандарты в области СПИ, на основе которых разрабатываются фактические стандарты для отдельных отраслей и процессов. Также развитие фактической стандартизации в области СПИ сегодня связано с формированием сводов правил и рекомендаций по применению, или их ещё называют, руководством по архитектуре (**Architecture Framework**). В этих руководствах описываются особенности практического использования апробированных архитектурных стилей и передового опыта разработки в различных отраслях хозяйственной деятельности.

Среди **фактических** стандартов можно выделить следующие:

- ISO/IEC TR 24748 — СПИ. Руководство по управлению ЖЦ. В контексте стандартов ISO/IEC 15288 и ISO/IEC 12207 определяет порядок использования моделей ЖЦ.
- ISO/IEC TR 19760:2003 — СИ. Руководство по применению ISO/IEC 15288. Содержит рекомендации по адаптации стандарта ISO/IEC 15288 к конкретным системам и проектам.
- ISO/IEC/IEEE 16326:2009 — СПИ. Процессы ЖЦ. Управление проектами. Взамен ISO/IEC TR 16326:1999 — Программная инженерия. Руководство по применению ISO/IEC 12207 при управлении проектами.
- ISO/IEC 16085:2006 — СПИ. Процессы ЖЦ. Управление рисками. Определяет процесс управления рисками в ЖЦ.

В России также существует две основные группы государственных стандартов, регламентирующих разработку ТЗ в сфере информационных технологий. Это стандарты серий ГОСТ 19 (единая система конструкторской документации (ЕСКД) и программной документации (ЕСПД)) и ГОСТ 34 (разработка систем автоматизации управления, проектирования, производства).

В 19 серии выделяется ГОСТ 19.201-78 «Техническое задание, требования к содержанию и оформлению» [1].

Согласно ГОСТ 19.201-78 типовое ТЗ должно включать следующие разделы:

- Введение.
- Основания для разработки.
- Назначение разработки.
- Требования к программе или программному изделию.
- Требования к программной документации.
- Техничко-экономические показатели.
- Стадии и этапы разработки.
- Порядок контроля и приёмки.
- Приложения.

Вторым стандартом является ГОСТ 34.602-89, который регламентирует структуру ТЗ на создание автоматизированной системы [2].

Согласно ГОСТ 34.602-89 типовое ТЗ должно включать следующие разделы:

- Общие сведения.

- Назначение и цели создания (развития) системы.
- Характеристика объектов автоматизации.
- Требования к системе.
- Состав и содержание работ по созданию системы.
- Порядок контроля и приёмки системы.
- Требования к составу и содержанию работ по подготовке объекта автоматизации к вводу системы в действие.
- Требования к документированию.
- Источники разработки.

Зарубежном ТЗ принято называть спецификацией требований, которая представляет собой специального вида документ для ПО, содержащий в себе информацию о том, как должна себя вести ИС, какие функции должна выполнять, какую нагрузку должна выдерживать и т.д.

Принято различать спецификацию системных требований (SyRS – System Requirements Specification) и спецификацию требований к ПО (SRS – Software requirements specification).

SyRS определяет технические требования для выбранной системы и удобства взаимодействия предполагаемой системы и человека, определяет высокоуровневые требования к системе с точки зрения предметной области, а также информацию об общей цели системы, ее целевой среде и ограничениях, допущениях и нефункциональных требованиях. SyRS может включать в себя концептуальные модели, спроектированные для иллюстрации содержания системы, сценариев использования, основных сущностей предметной области, данных, информации и рабочих процессов. В этом контексте SyRS представляет собой аналог отечественного ТЗ по ГОСТ 34.

SRS – это спецификация требований для определённого программного изделия, программы или набора программ (продукт), которые выполняют определенные функции в конкретном окружении. SRS включает ряд пользовательских сценариев (usecases), которые описывают все варианты взаимодействия между пользователями и ПО.

Пользовательские сценарии являются средством представления функциональных требований. В дополнение к пользовательским сценариям, спецификация также содержит нефункциональные требования, которые налагают ограничения на дизайн или реализацию (такие как требования скорости вычислений, объёма памяти, стандарты качества, или проектные ограничения).

В данном контексте SRS представляет собой аналог отечественного ТЗ по ГОСТ 19.

Главным отличием отечественных ТЗ по ГОСТ 19 и 34 является то, что они определяют требования к проведению работ, к обеспечению работ и испытаний. Обычно, даётся ссылка на документ «программа и методика испытаний».

Таким образом, ТЗ по ГОСТ 34 дополнительно определяет, что должно быть сделано (ключевые признаки успеха проекта), в какие сроки и каковы промежуточные этапы (точки контроля). Также из ТЗ по ГОСТ 34 принято выделять частные ТЗ (ЧТЗ), на более мелкие работы.

### **Выбор средств программной реализации веб-сервиса**

Подход к автоматизации формирования ТЗ по ГОСТ серий 19 и 34 предполагает повторное использование накопленных знаний технических писателей (пользователей) с помощью веб-сервиса, который автоматически генерирует фрагменты контента для разделов ТЗ согласно ранее написанным и предварительно оценённым пользователями фрагментам похожих по тегам ТЗ.

Для программной реализации веб-сервиса выбраны следующие средства:

- Язык программирования TypeScript.

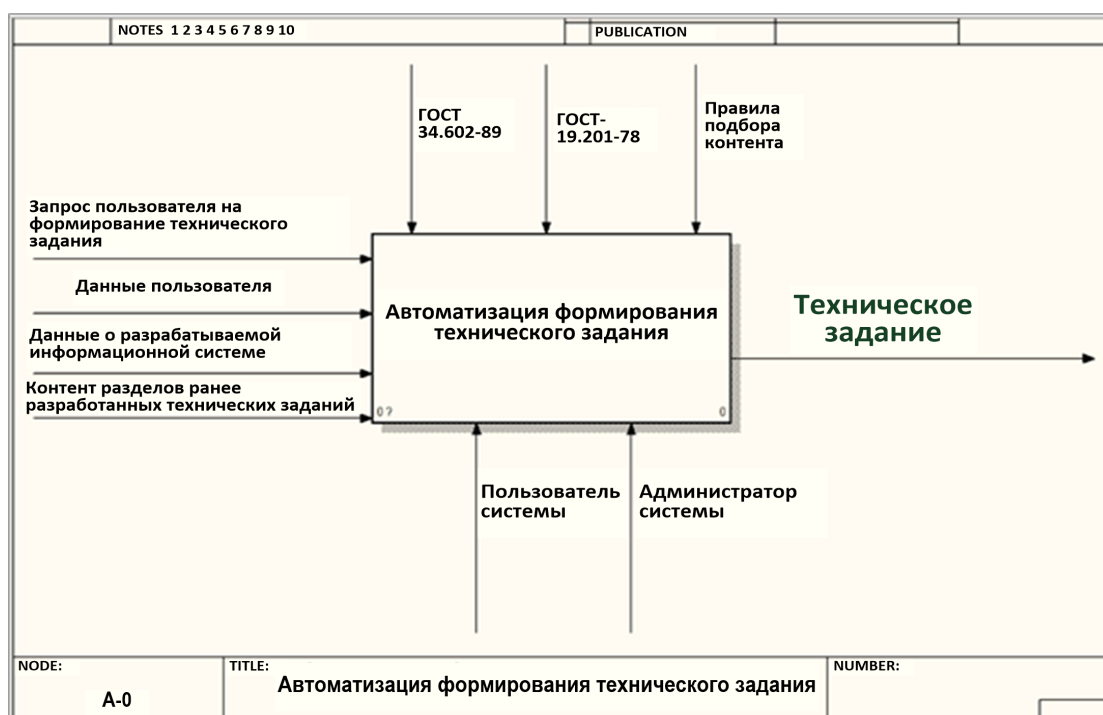
- Jet Brains Web Storm – интегрированная среда разработки на Java Script, CSS и HTML от компании Jet Brains.
- Node.js – платформа для построения эффективных и масштабируемых приложений
- React.js – фреймворк для создания модульных пользовательских интерфейсов.
- Postgre SQL – свободная объектно-реляционная система управления базами данных.

Веб-сервис должен обеспечивать выполнение следующих основных функций:

- регистрация и авторизация пользователей;
- создание и редактирование ТЗ по ГОСТ-19.201-78 с учетом рекомендаций процессного ГОСТ Р ИСО/МЭК 12207-2010 «Информационная технология. Системная и программная инженерия. Процессы жизненного цикла ПО»;
- создание и редактирование ТЗ по ГОСТ-34.602-89;
- накопление тегов ТЗ, фрагментов контента и их оценок;
- формирование рекомендаций содержания разделов ТЗ на основе тегов и оценок фрагментов контента;
- выгрузка ТЗ в виде PDF документа.

### Проектирование архитектуры и алгоритмов функционирования веб-сервиса

В процессе проектирования была составлена контекстная диаграмма [5] веб-сервиса, отражающая основной процесс по автоматизации формирования ТЗ (рис. 1).



**Рис. 1. Контекстная диаграмма верхнего уровня IDEF0**

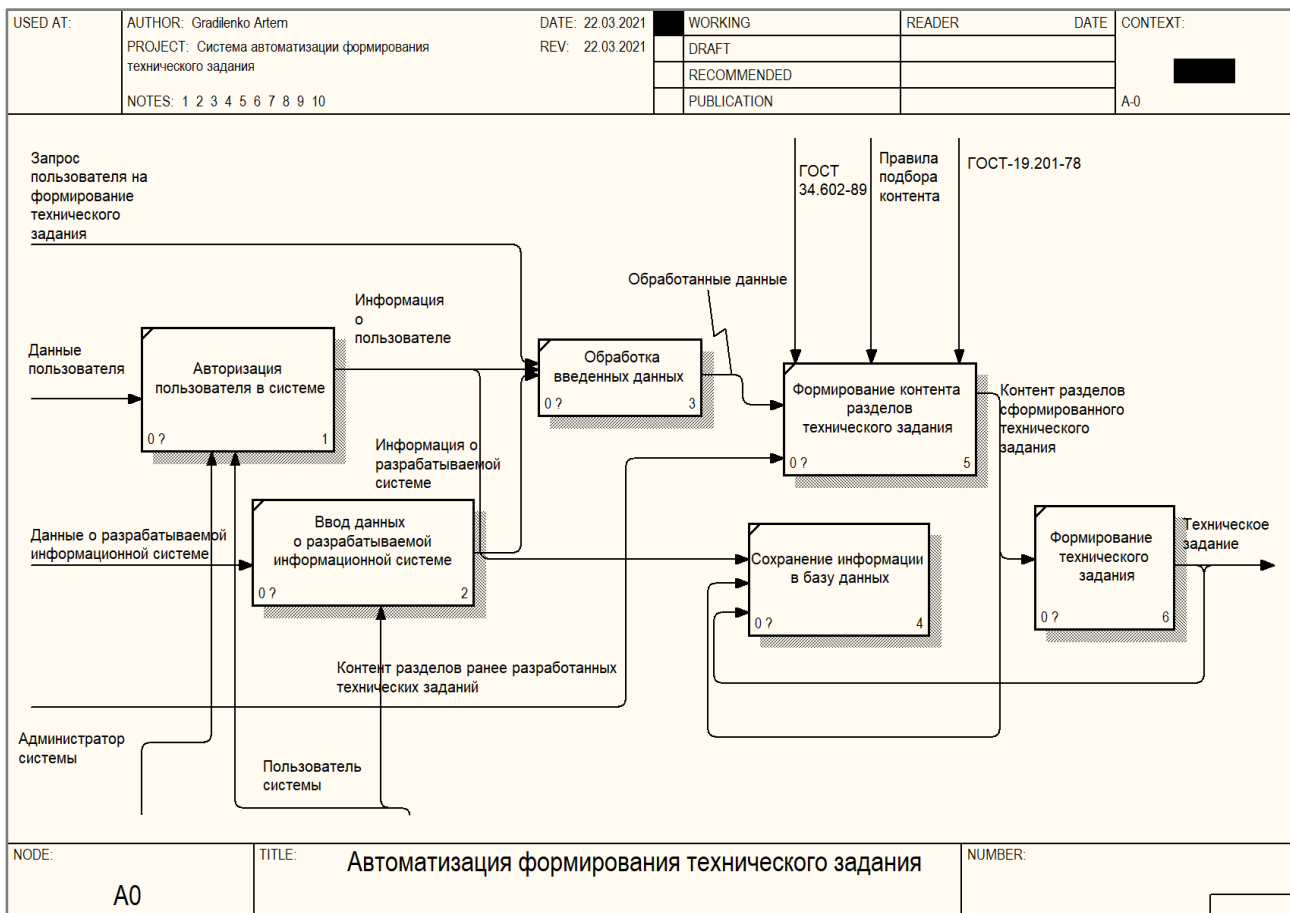
На вход поступают данные пользователя, данные о разрабатываемой ИС и сведения о ранее разработанных ТЗ.

С веб-сервисом взаимодействуют пользователь и администратор веб-сервиса. Управляющими инструментами являются рассмотренные ранее ГОСТ и правила подбора текстового контента.

На выходе пользователь получает заполненный шаблон ТЗ.

В архитектуре веб-сервиса выделяются следующие блоки: блок авторизации, блок ввода данных, блок обработки данных, блок формирования контента разделов ТЗ, блок сохранения в базу данных, блок формирования ТЗ.

На рис. 2 представлена диаграмма декомпозиции веб-системы в нотации IDEF0.



**Рис. 2. Диаграмма декомпозиции системы**

Назначение блоков:

Блок авторизации – блок, отвечающий за авторизацию пользователей в системе, регистрацию новых пользователей.

Блок ввода данных о разрабатываемой системе – блок, отвечающий за ввод в систему данных о ИС, для которой будет разрабатываться ТЗ.

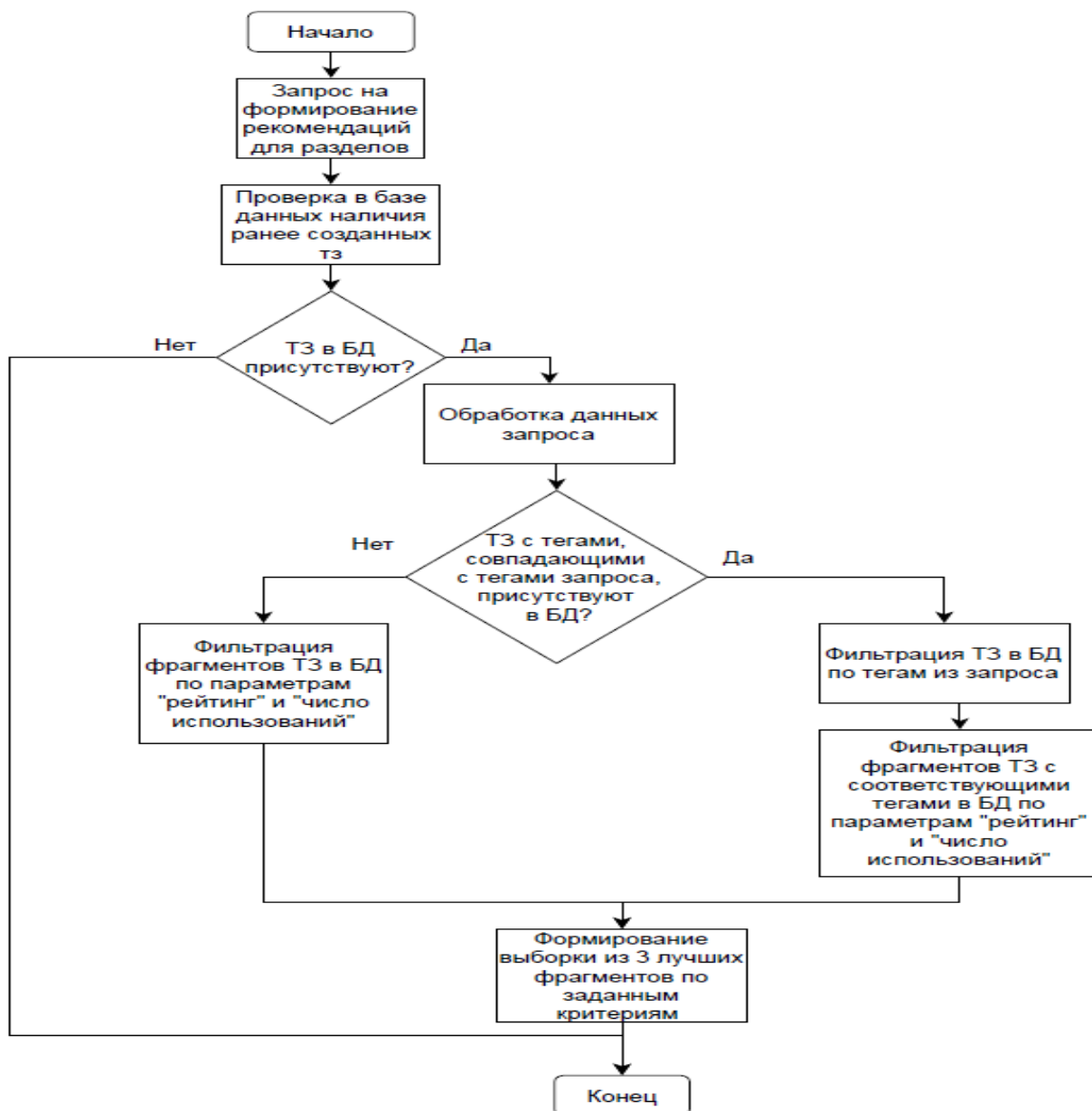
Блок обработки данных – вспомогательный блок, отвечающий за обработку данных, поступающих от пользователя за время использования системы.

Блок формирования контента разделов технического задания – блок, отвечающий за генерацию рекомендаций для содержания разделов разрабатываемого ТЗ.

Блок сохранения в базу данных – блок, отвечающий за сбор информации о технических заданиях, их содержанием и пользователям системы.

Блок формирования технического задания – блок, отвечающий за создание технического задания и генерацию технического задания в виде текстового документа формата PDF.

На рис.3 представлена блок-схема алгоритма генерации фрагментов контента для заполнения соответствующего раздела ТЗ.



**Рис.3. Блок-схема алгоритма генерации фрагментов контента для заполнения соответствующего раздела ТЗ**

Блок-схема алгоритма генерации фрагментов контента для заполнения соответствующего раздела ТЗ включает в себя следующие функции:

- *запрос на формирование рекомендации для разделов* – данные о создаваемом пользователем ТЗ (тематика ТЗ, выбранный ГОСТ) поступают в блок, для их дальнейшей обработки;
- *проверка в базе данных наличия ранее созданных ТЗ* – система проверяет наличие в базе данных ранее созданных ТЗ;
- *обработка данных запроса* – система считывает данные запроса на формирование рекомендаций и выделяет из структуры необходимые данные, заносит их в память для дальнейшей передачи;
- *фильтрация фрагментов ТЗ в БД[7] по параметрам «рейтинг» и «число использований»* – система осуществляет фильтрацию ранее сохранённых фрагментов ТЗ в базе данных по параметрам «рейтинг» и «число использований» по убыванию, т.е. в начале коллекции после фильтрации оказываются фрагменты с

наибольшим суммарным показателем по двум параметрам, а в конце – с наименьшим;

- *фильтрация ТЗ в БД по тегам из запроса* – система осуществляет фильтрацию ранее сохранённых в базе данных ТЗ с соответствующим ГОСТ по тегам, указанным в запросе на формирование рекомендаций, т.е. система формирует выборку ТЗ, теги которых совпадают с тегами из запроса;
- *фильтрация фрагментов ТЗ с соответствующими тегами в БД по параметрам «рейтинг» и «число использований»* – система производит фильтрацию выборки из предыдущего пункта по параметрам «рейтинг» и «число использований» по убыванию, т.е. в начале коллекции после фильтрации оказываются фрагменты с наибольшим суммарным показателем по двум параметрам, а в конце – с наименьшим;
- *формирование выборки из 3 лучших фрагментов по заданным критериям* – система формирует выборку из трёх лучших содержаний фрагментов ТЗ из коллекции, сформированной и отфильтрованной на предыдущем шаге.

Алгоритм генерации рекомендаций для содержимого раздела ТЗ основан на работе со значениями фрагментов, ранее сохранённых в базе данных ТЗ. В сути алгоритма заложено следующее: в то время, как пользователь начинает заполнять содержимое ТЗ, блок получает запрос, содержащий справочные данные, указанные пользователем, при создании ТЗ.

Далее, система проверяет наличие ранее сохранённых ТЗ в БД. Если таковые отсутствуют, алгоритм заканчивает работу. Если в БД присутствуют ранее созданные ТЗ, система обрабатывает запрос, выделяя из него следующие параметры [8]: ГОСТ и теги.

Далее, система осуществляет фильтрацию ранее сохранённых ТЗ, сохраняя в выборку ТЗ с совпадающим ГОСТ и тегами, если таковые присутствуют в БД. В случае отсутствия ТЗ с необходимыми тегами, система обеспечивает фильтрацию всех находящихся в БД ТЗ, фильтруя их по параметрам «рейтинг» и «число использований» по убыванию, т.е. в начале коллекции после фильтрации оказываются фрагменты с наибольшим суммарным показателем по двум параметрам, а в конце – с наименьшим. И отправляет пользователю 3 фрагмента, обладающих наибольшим суммарным показателем.

В случае присутствия ТЗ в БД с совпадающими тегами, система заносит такие фрагменты в отдельную выборку. Далее, система производит фильтрацию выборки из предыдущего шага по параметрам «рейтинг» и «число использований» по убыванию, т.е. в начале коллекции после фильтрации, также как и до этого, оказываются фрагменты контента с наибольшим суммарным показателем по двум параметрам, а в конце – с наименьшим.

В итоге, пользователь получает 3 фрагмента, обладающих наибольшим суммарным показателем, из которых он может выбрать один, чтобы заполнить своё ТЗ и оценить его по 5 бальной шкале. Если фрагмент не подходит, то пользователь пишет свой собственный текст.

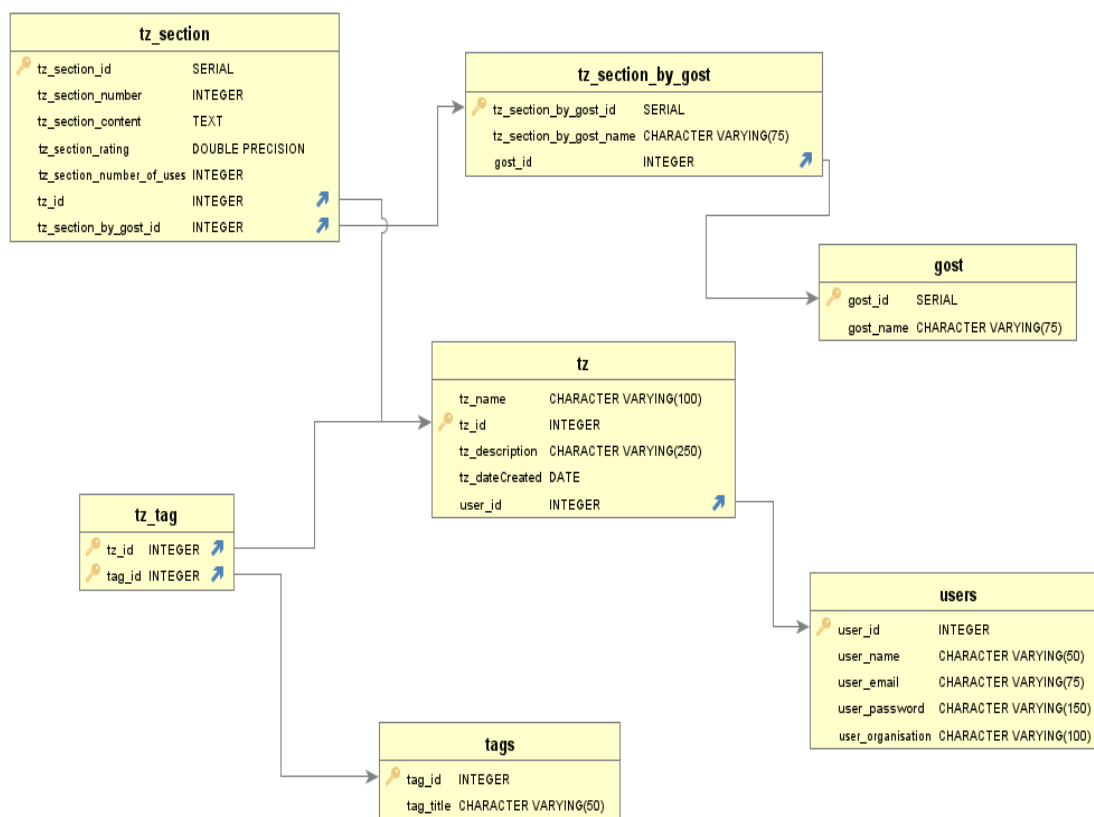
### **Проектирование схемы базы данных веб-сервиса**

На рис. 4 представлена разработанная схема базы данных веб-сервиса.

База данных состоит из 7 таблиц:

- users – таблица, содержащая информацию о пользователях системы;
- tz – таблица, содержащая информацию о технических заданиях, хранящихся в системе;
- gost – таблица, содержащая информацию о ГОСТах, используемых в системе;
- tags – таблица, содержащая информацию о тегах, придаваемых ТЗ;
- tz-tag – вспомогательная таблица, содержащая информацию о тегах, принадлежащих конкретному экземпляру из таблицы tz;
- tz\_section – таблица, содержащая информацию о разделах ТЗ и их содержании;

- tz\_section\_by\_gost– таблица, содержащая информацию о разделах ТЗ согласно ГОСТ.



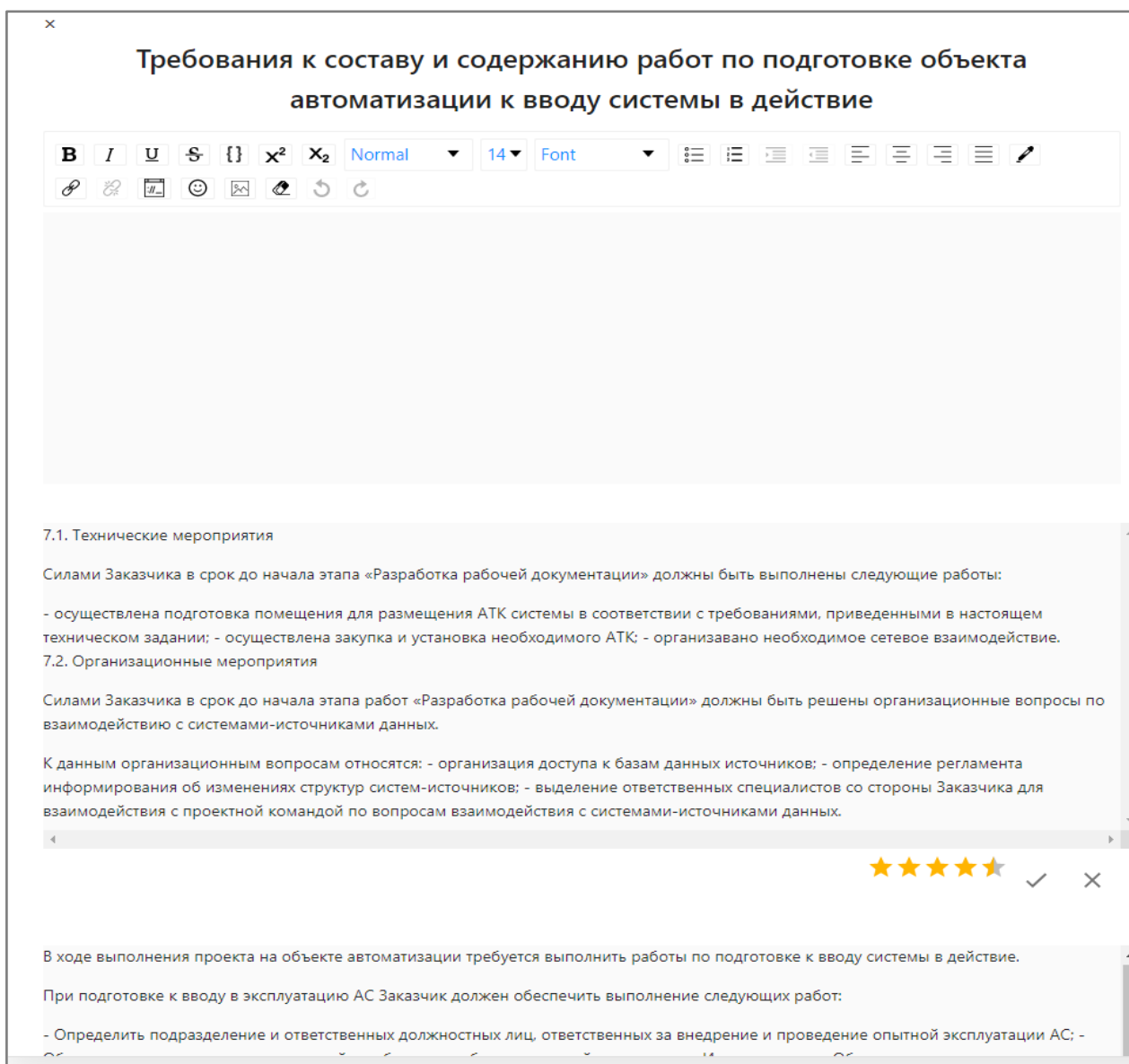
**Рис.4. Схема базы данных веб-сервиса**

### **Пример работы с разработанным веб-сервисом**

На рис. 5 представлено окно веб-сервиса для редактирования раздела создаваемого вновь типового ТЗ.

В верхней части экранной формы отображается: заголовок раздела, WYSIWYG [8] редактор, в котором пользователь может заполнить содержимое раздела ТЗ при помощи всех элементов, присущих обычным текстовым редакторам. Ниже располагается блок выбора из рекомендаций, в котором отображаются 3 сгенерированные автоматически рекомендации. Под каждой рекомендацией расположен ряд, включающий кнопки для выставления рейтинга, кнопки «подтвердить» и «отклонить». При нажатии на кнопку «подтвердить», содержимое рекомендации переносится в поле WYSIWYG редактора. При нажатии кнопки «отклонить», данная рекомендация удаляется из блока рекомендаций.

Ниже блока рекомендаций находится кнопка «Сохранить информацию», нажатие которой сохраняет информацию о содержимом раздела, а также кнопка «Очистить поле ввода», нажатие на которую очищает поле ввода WYSIWYG редактора.



**Рис. 5. Окно веб-сервиса для редактирования раздела создаваемого ТЗ**

### **Заключение**

Веб-сервис может быть использован предприятиями, сферой деятельности которых является проектирование и разработка ИС для автоматизации процесса формирования ТЗ, что позволит сократить затрачиваемые временные ресурсы и повторно использовать накопленные знания технических писателей. Последующая доработка веб сервиса позволит использовать его не только для создания ТЗ, но и для процесса управления требованиями.

По результатам работы было получено свидетельство об официальной регистрации программы для ЭВМ № 2021663264 от 13.08.2021 г.



### Библиографический список

1. **ГОСТ 19.201-78.** Единая система программной документации. Техническое задание. Требования к содержанию и оформлению: Межгосударственный стандарт: дата введения 1980-01-01 / Государственный комитет СССР по стандартам. – Изд. официальное. – Москва:Стандартинформ, 2010. – 3 с.
2. **ГОСТ 34.602-89.** Информационная технология. Комплекс стандартов на автоматизированные системы. Техническое задание на создание автоматизированной системы: дата введения 1990-01-01 / Государственный комитет СССР по стандартам. – Изд. официальное. – Москва :Стандартинформ, 2009. – 12 с.
3. **ГОСТ Р ИСО/МЭК 12207 - 2010.** Информационная технология. Системная и программная инженерия. Процессы жизненного цикла программных средств. Национальный стандарт: дата введения 2010-11-30/Технический комитет по стандартизации ТК22 «Информационные технологии».- Изд. официальное. – Москва :Стандартинформ, 2011. – 105 с.
4. Использование инструментов IBM Rational/Telelogic в рамках управления жизненным циклом разработки приложений. IBM Rational/Telelogic **DOORS®/Analyst**. [Электронный ресурс]:- Режим доступа: [www.interface.ru](http://www.interface.ru)(дата обращения: 11.05.2022).
5. **Хаф, Л.** Методология разработки программного обеспечения: в 3-х ч.- Ч.3: RationalUnifiedProcess: [Электронный ресурс]:- Режим доступа: [http://www.lib.csu.ru/dl/bases/prg/kompress/articles/2004\\_01\\_rupIntro/index.html](http://www.lib.csu.ru/dl/bases/prg/kompress/articles/2004_01_rupIntro/index.html)(дата обращения: 12.03.2022).
6. **Орлов, С.А.** Технологии разработки программного обеспечения: учеб. пособие / С.А. Орлов. - 2-е изд. - СПб.: Питер, 2003. - 480 с.
7. **Одинцов, И.О.** Профессиональное программирование. Системный подход / И.О. Одинцов. - 2-е изд., перераб. и доп. - СПб.: БХВ-Петербург, 2004. - 624 с.
8. **Фримен, А.** ASP.NET 4.5 с примерами на C# 5.0 для профессионалов, 5-е издание = Pro ASP.NET 4.5 in C#, 5th Edition. — М.: «Вильямс», 2014. — 1120 с. — ISBN 978-5-8459-1878-9.

**ИССЛЕДОВАНИЕ И РАЗРАБОТКА МЕТОДОВ ПО УСОВЕРШЕНСТВОВАНИЮ  
ПЕРЕДАЧИ ШИРОКОВЕЩАТЕЛЬНЫХ СООБЩЕНИЙ В СЕТЯХ СОТОВОЙ СВЯЗИ 2G-4G  
ПОКОЛЕНИЙ НА ПРИМЕРЕ СИСТЕМЫ-ПРОТОТИПА**

**М.В. Ерютов**

ORCID: 0000-0002-8673-9601 e-mail: m.v.erytow@yandex.ru

Нижегородский государственный технический университет им. Р.Е.Алексева

В докладе производится анализ прототипа сети GSM при передаче широковещательных сообщений. Анализ производится с помощью проекта OSMOCOM, имплементирующего стек протоколов GSM. Прототип сети построен с помощью данного программного обеспечения на персональном компьютере и мобильного устройства, выступающего в качестве базовой станции. С помощью прототипа организована сеть сотовой связи стандарта 2G. Предложен вариант использования каналов CellBroadcastChannel (CBCH) для передачи широковещательных сообщений в 26 каналах трафика. Последний кадр из мультикадров каналов трафика может быть настроен на основную частоту базовой станции, нулевой временной интервал. Такая настройка может привести к тому, что положение последнего кадра мультикадра TCH смещается относительно мультикадра управления BCCH ровно на один кадр каждые 240 мс (что равно двум TDMA кадрам трафика). При переходе на основную частоту базовой станции и нулевой таймслот в интервале 26 кадра мобильная станция определяет логический канал, и в случае, если это канал CBCH или если номер кадра в мультикадре каналов управления с 32 по 35, она должна получить информацию этого канала.

**Ключевые слова:** CellBroadcast, широковещательная передача, сотовая связь, протоколы сотовой связи, базовые станции GSM, 2G, 3G, 4G, OSMOCOM.

**RESEARCH AND DEVELOPMENT OF METHODS FOR IMPROVING THE  
TRANSMISSION OF BROADCAST MESSAGES IN CELLULAR NETWORKS OF 2G-4G  
GENERATIONS ON THE EXAMPLE OF A PROTOTYPE SYSTEM**

**M.V. Eryutov**

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

**Purpose:** The author provides a method for transmitting broadcast messages in traffic channels.

**Design / methodology / approach:** The article describes the possibilities of the OSMOCOM project, the process of creating a 2G prototype system that allows you to organize the connection of two mobile phones through a base station. Author offers A variant of using Cell Broadcast Channels (CBCH) is proposed for transmission of broadcast messages in 26 traffic channels. The last frame of the traffic channel multiframe may be tuned to the base station fundamental frequency, slot zero. Such a setting may cause the position of the last frame of the TCH multiframe to shift relative to the BCCH control multiframe by exactly one frame every 240 ms (equal to two TDMA traffic frames). When switching to the base station fundamental frequency and zero timeslot in frame interval 26, the mobile station determines the logical channel, and if it is a CBCH channel or if the frame number in the multiframe of control channels is from 32 to 35, then it should receive information of this channel.

**Findings:** In the paper, an optimal variant of broadcast messages transmission is found, which ensures the delivery of messages to all subscribers.

**Research limitations/implications:** Research is limited to the use of mobile devices with Calypso chipsets only. But the results can be embedded in real equipment such as base stations.

**Originality/value:** The value of the entire article is the new method of broadcasting messages.

**Key words:** Cell Broadcast, cellular protocols, base stations GSM, 2G, 3G, 4G, OSMOCOM.

## Введение

На сегодняшний день в системах сотовой связи стандарта GSM нет гарантии получения абонентом широковещательных сообщений в случае чрезвычайной ситуации и применении технологии Cell Broadcast (CB). Это происходит, во-первых, потому что в случае, когда мобильная станция (МС) использует канал трафика и одновременно с этим осуществляется рассылка широковещательных сообщений, то в силу возможностей технологии CB такие сообщения будут пропущены МС. Во-вторых, пользователь может отключить прием широковещательных сообщений. Первая проблема решается установкой периода повторной передачи широковещательных сообщений. Но это ведет к тому, что нагрузка на каналы управления увеличивается, что является также проблемой. Вторую проблему, к сожалению, разрешить невозможно, поскольку технология CB предоставляет эту возможность отключать прием широковещательных сообщений. Отключение этой возможности значило бы для многих нарушение их права выбора. Поэтому необходимо направить силы на решение первой проблемы. Но, поскольку мы не располагаем настоящим оборудованием для анализа сети стандарта GSM и Cell Broadcast, в частности, то необходимо использовать сторонние разработки. Решением этой проблемы может оказаться применение открытого программного обеспечения проекта Osmocom.

## Анализ системы Cell Broadcast

Cell Broadcast (CB) — это технология одновременной отправки сообщений нескольким пользователям мобильных телефонов в определенной области. Технология представлена комитетом GSM ETSI и 3GPP и является частью стандартов 2G, 3G, 4G LTE (телекоммуникации) и 5G.

Передачу широковещательного сообщения организует клиент, например, правительство, службы спасения, метеослужбы или коммерческие организации, через систему Cell Broadcast Entity (CBE) и сервер Cell Broadcast Center (CBC). Через CBC осуществляется передача сообщения к контроллерам базовых станций по протоколам Cell Broadcast Service Protocol (CBSP), Service Area Broadcast Protocol (SABP), Service Broadcast Application Part (SBc-AP) в 2G, 3G и 4G соответственно.

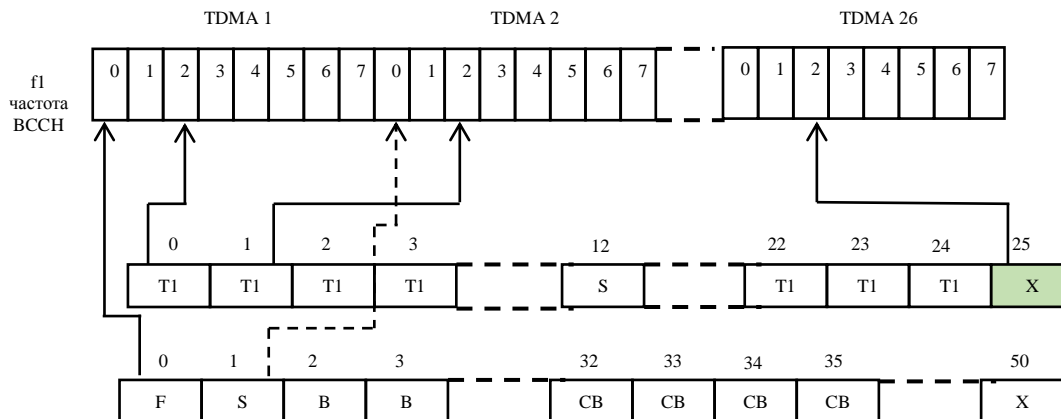
Для передачи широковещательных сообщений в сетях 2G используется логический канал Cell Broadcast Channel (CBCH), который принадлежит группе широковещательных каналов Broadcast Channel (BCH). Типичный сценарий использования канала CBCH – использование совместно с каналом Standalone Dedicated Control Channel (SDCCH/8), при этом передача осуществляется во втором слоте TS1 или в первом слоте TS0. Пример одной из таких конфигураций показан на рис.1 (каналы CBCH отмечены в мультикадре каналов управления в кадрах с 32 по 35).

Cell Broadcast-сообщение имеет следующую структуру:

- Сообщение имеет максимум 15 страниц.
- Каждая страница представляет собой 82 байта данных, в результате чего получается 93 символа в 7-битном алфавите GSM по умолчанию[3].

Информация по каналу CBCH передается в Normal Burst (NB) пакетах. Перед отправкой сообщения данные формируются в блоки по 184 бита в каждом. Затем эти биты поступают в кодер, выполняющий помехоустойчивое кодирование блочным циклическим кодом (224, 184). К полученному коду добавляются хвостовые биты в количестве 4 штук. Затем последовательность бит поступает в кодер, выполняющий свёрточное кодирование со скоростью  $r=1/2$  (456,228). Таким образом, для передачи блока из 184 бит используется четыре NB пакета. Поскольку  $184/8=23$  байта, а для передачи одного широковещательного сообщения или одной страницы необходимо передать 4 блока сообщения по 22 байта в каждом и плюс один байт для определения типа блока, передаваемого контента и номера блока; то, в общем,

необходимо передать  $4 \times 4 = 16$  пакетов NB. Скорость передачи широковещательных сообщений равна 0.782 кбит/с.



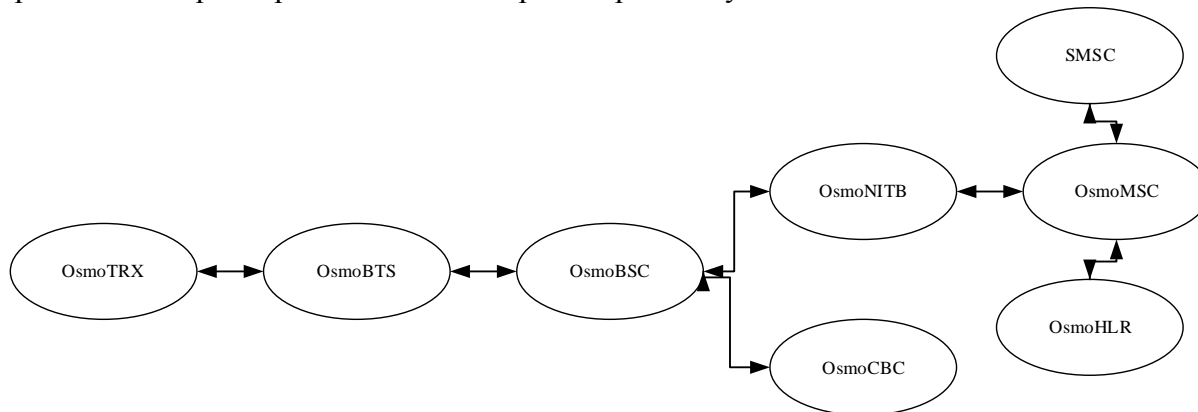
**Рис. 1. Пример конфигурации мультикадра каналов управления**

В сетях сотовой связи стандарта 3G для передачи широковещательных сообщений используется логический канал (Common Traffic Channel), который отображается на канал транспортного уровня FACCH (Forward Access Channel), который отвечает за передачу сигнальной, широковещательной или многоадресной информации к мобильным станциям.

В сетях сотовой связи стандарта 4G отказались от использования технологии Cell Broadcast [4,7]. Вместо неё используется система оповещения Public Warning System (PWS). Широковещательные сообщения передаются по логическому каналу Broadcast Control Channel (BCCH). Затем информация передается каналу транспортного уровня Broadcast Channel (BCH) и дальше каналу Physical Broadcast Channel (PBCH).

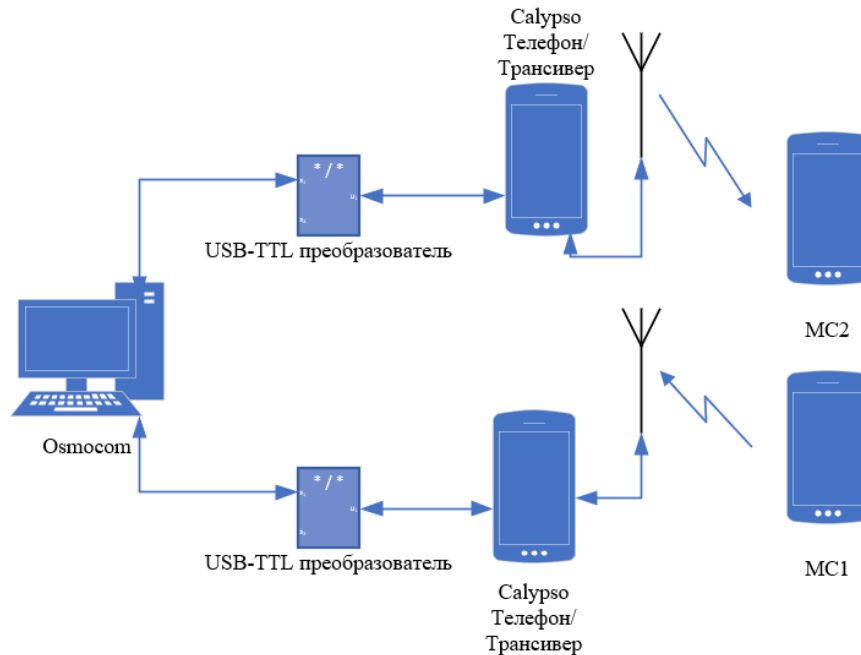
### Анализ широковещательной передачи на прототипе сотовой сети

Для проведения анализа широковещательной передачи воспользуемся проектом Osmocom. Данный проект имплементирует (реализует) стеки протоколов сетей сотовой связи стандартов 2G, 3G и 4G. Это означает возможность реализации всех компонентов сетей сотовой связи на уровне программ. Для наглядности на рис.2 иллюстрируется реализации и взаимодействия компонентов Osmocom для сети 2G. Здесь Osmo NITB является аналогом подсистемы базовой станции, но на программном уровне. Взаимодействие с программными компонентами такой сети происходит через терминал компьютера по протоколу telnet.



**Рис. 2. Структура проекта Osmocom для сети 2G**

В качестве устройства, выполняющего функцию базовой станции, выступают телефоны на базе чипсета Calypso [5]. Через 2.5 мм разъем телефона осуществляется их подключение к компьютеру через USB-TTL конвертер. Кроме телефонов могут использоваться также SDR-приемопередатчики [6]. Структура сети-прототипа, построенной на основе телефонов на базе чипсетов Calypso, имеет вид показанный на рис.3. Такая схема позволяет организовать сеть стандарта GSM с возможностью осуществлять звонки, отправлять SMS сообщения и широковещательные сообщения. Существует и другая схема, в которой используется один мобильный телефон. В такой схеме уже могут использоваться другие программы, которые используются для просмотра управляющей информации, сбора, хранения и анализа трафика.



**Рис.3. Структурная схема сети-прототипа на основе проекта Osmocom**

Для настройки такой сети используются всего два файла: конфигурационный файл настройки контроллера базовой станции, называемый `osmo-bsc.cfg`, и конфигурационный файл настройки базовой станции, называемый `osmo-bts.cfg`. Для настройки широковещательной передачи в сети стандарта 2G применяются следующие команды в файле `osmo-bsc.cfg`:

```

Network
handover 0 # Отключение хэндовера
bts 0# номер базовой станции;
trx 0 # номер трансивера;
arfcn 10 # номер частоты вещания (GSM900: UL 892, DL 937)
timeslot 0 # номерслота;
phys_chan_configCCCH+SDCCH4+CBCH # конфигурацияканала.
timeslot 1
phys_chan_configSDCCH8
timeslot2
phys_chan_configTCH/F_TCH
timeslot3

```

```
phys_chan_configTCH/F_TCH
timeslot4
phys_chan_configTCH/F_TCH
timeslot5
phys_chan_configTCH/F_TCH
timeslot6
phys_chan_configTCH/F_TCH
timeslot7
phys_chan_config TCH/F_TCH
```

Основные настройки файла конфигурации базовой станции имеют вид:

```
phy 0
osmotrx ip 127.0.0.1
osmotrx fn-advance 30
osmotrx rts-advance 5
instance 0bts 0
band GSM900
trx 0
rxgain 0
power 0
```

Здесь представлены основные настройки позволяющие задать частотный диапазон, число частотных каналов, конфигурацию логических каналов.

В настоящее время ведется развитие проекта Osmocom. Помимо стандарта 2G была добавлена поддержка стандартов 3G и 4G. Однако пока не реализованы некоторые компоненты для создания полноценных сетей 3G или 4G. В частности, пока нет возможности реализовать ширококвещательную передачу. Таким образом, сочетание программного обеспечения и обычного мобильного телефона позволяет организовать небольшую, но имеющую все функции полноценной, сеть сотовой связи.

#### **Анализ возможности усовершенствования передачи ширококвещательных сообщений**

В мультикадре каналов трафика сети 2G используется 26 кадров, из которых 24 используется для передачи голоса. Последний кадр (на рис.1 выделен цветом) используется для того, чтобы принимать информацию ширококвещательных каналов Synchronization Channel (SCH) и Frequency Correction Channel (FCCN) [1,2], которые передают информацию о кадровой и частотной синхронизации. Прием информации каналов SCH и FCCN возможен благодаря разной длине мультикадров каналов управления и трафика. Допустим мобильной станции уже назначен слот и используется канал трафика для передачи голоса. После передачи или приёма кадра с номером 24 мобильная станция переключается в режим приема ширококвещательной информации, в этом режиме MS ищет пакеты каналов SCH и FCCN. Поскольку число кадров мультикадра каналов управления не кратно числу кадров каналов трафика, то через каждые 120 мс, соответствующие длительности 26 TDMA кадров канала управления, происходит смещение относительно мультикадра каналов управления ровно на один кадр. Выше сказанное иллюстрирует рис.4.

Кроме того, существует несколько режимов передачи голоса. Интерес представляет полускоростной режим. Мобильная станция передает информацию по каналам трафика в Normal Burst (NB) пакетах, которые состоят из двух полезных нагрузок размером по 57 бит. В

полноскоростном режиме для передачи речи задействованы обе нагрузки. В полускоростном режиме одна из нагрузок может быть отдана или другому абоненту, или может использоваться для передачи сигнальной информации по каналу Fast Associated Control Channel (FACCH). Необходимость в использовании канала FACCH может возникнуть, когда требуется передать сигнальную или управляющую информацию если телефон находится в режиме передачи голосового трафика. Этой управляющей информацией может быть команда разрыва голосового соединения, хэндовер или другая сигнальная информация.

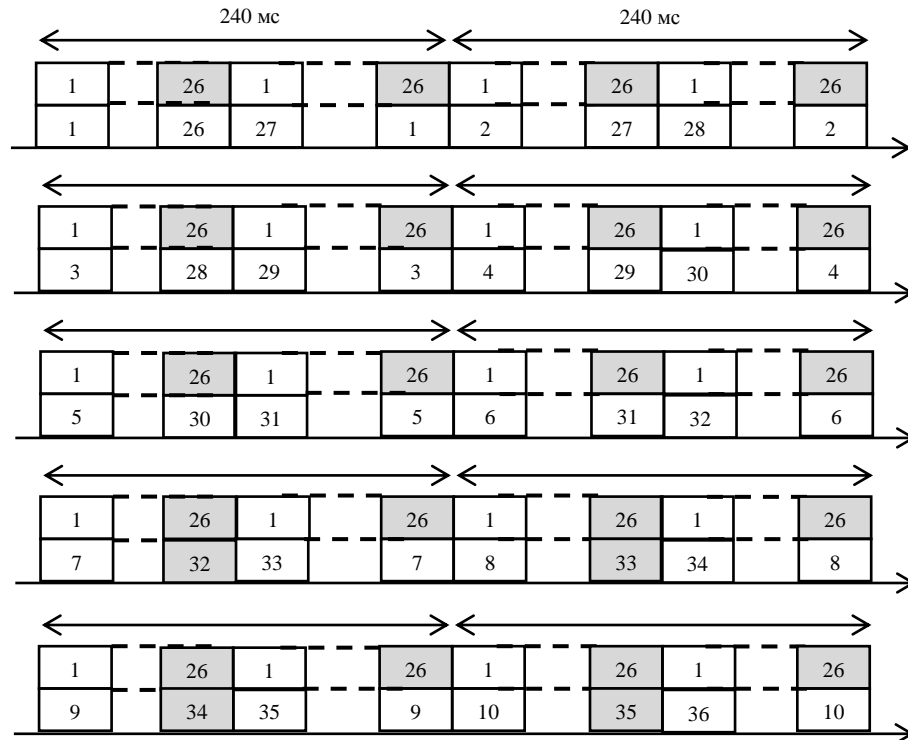


Рис. 4. Применение 26 кадра для приема информации канала СВСН

На рис.5 представлена структура пакета NB. Здесь для определения передачи информации канала FACCH используются поля “Флаг”. При совмещенной передаче голоса и сигнальной информации для одного бита полей “Флаг” устанавливается значение 1, а для другого поля значение 0.



Рис.5. Структура пакета Normal Burst

Таким образом, можно предложить два способа модернизации приема широковещательных сообщений:

1) Прием широковещательных сообщений в 26 кадре мультикадра каналов трафика.

2) Использование канала FACH для разрыва соединения или уведомления для перехода в режим приема широковещательных сообщений.

Первый способ эффективен тем, что не прерывает голосовую связь. Определим время доставки широковещательного сообщения и скорость передачи, зная, что размер сообщения равен 88 байт, а для его передачи потребуется 16 NB пакета, то есть MS требуется четыре раз переключиться на широковещательный канал для получения всего сообщения. В случае, если используется конфигурация, показанная на рис.1, то время передачи равно  $2 * 51 * 235.365 = 24007.23$  мс (умножение на 2 соответствует тому, что за время равное 51 мультикадру каналов управления принимается 8 широковещательных NB пакетов). Скорость передачи при таком способе равна  $\frac{23*4*8}{2*51*51*8*576.9} = 0.031$  кбит/с.

Второй способ позволяет гарантированно доставить широковещательное сообщение, но при этом нагружает сотовую сеть и прерывает связь.

Рассматривая варианты развития широковещательной передачи широковещательной информации, стоит упомянуть не только о разработанных стандартах для передачи необходимой информации по широковещательным каналам связи таких как Cell Broadcast или Public Warning System. Поскольку в последнее время возрастает доля интернет-телефонии [8], то одним из вариантов передачи широковещательных сообщений может стать специально разработанное для этого приложение, которое устанавливается на мобильное устройство пользователя, собирает информацию о местоположении и передает её поставщику услуг приложений. Это позволяет поставщику приложений отправлять приложение предупреждающие сообщения через Интернет. Задача технологии состоит в том, чтобы привлечь достаточное количество людей для установки приложения. В этом состоит главная проблема такого подхода. Также другой проблемой оказывается повышение нагрузки на сеть в случае ЧС, когда другие подходы, использование CBS или PWS, не нагружают сеть. Достоинствами приложений для широковещательной передачи являются отсутствие сотрудничества с операторами сотовой связи и поддержка всех устройств, имеющих доступ в Интернет.

Ещё один вариант для доставки экстренного оповещения подавляющему большинству мобильных пользователей, не зависящий от мобильного устройства или приложения заключается в использовании апплета SIM-карты. Оповещения могут передаваться в апплет SIM-карты с помощью Cell Broadcast (загрузка широковещательных данных CBdd – cell), SMS(SMSdd) или IP (по протоколу VIP –Bearer Independent Protocol). Кроме того, данный способ может помочь при реализации приема широковещательных сообщений в 26 кадре канала трафика.

Кроме того, нельзя исключать возможность отключения приема и передачи всей трафиковой информации в случае чрезвычайной ситуации для незамедлительной трансляции широковещательных сообщений. Такой вариант уже был предложен [9], но остается вопрос гуманности этого подхода по отношению к пользователям сотовой связи. Разумнее было бы использовать уже существующие каналы сигнализации в разных стандартах сотовой связи, например, канал FACH в сети стандарта 2G, для кратковременного отключения всех трафиковых каналов и передачи широковещательных сообщений. Сценарий данного подхода может быть следующим: в случае ЧС учреждение (службы спасения, правительство и др.) определяет размер области, в которую нужно транслировать широковещательные сообщения; затем операторы сотовой связи через контроллеры базовых станций по сигнальным каналам на короткое время прерывают голосовую связь и по широковещательным каналам начинается



трансляция широковещательных сообщений; после прекращения вещания восстанавливается передача голосового трафика. Для определения области вещания необходимы программные средства и оборудование, которые были рассмотрены выше (СВЕ, СВС), а также сотрудничество с операторами сотовой связи для получения идентификаторов базовых станций.

### **Выводы**

Описана реализация сети сотовой связи стандарта 2Gc помощью проекта Osmocom. Основными компонентами при создании некоммерческой сети GSM на базе Osmocom являются программные реализации подсистемы базовой станции и совместимое оборудование (SDR-приемопередатчик или мобильный телефон на базе чипсета Calypso), выполняющее роль базовой станции. Описаны возможные варианты модернизации передачи широковещательных данных в сотовых сетях наземной радиосвязи различных поколений. В сетях стандарта GSM найден вариант приема широковещательных сообщений при голосовом вызове. Механизм приема сообщения схож с механизмом приема информации каналов кадровой синхронизации (SCH) и коррекции частоты (FCCH), при котором мобильная станция на интервале приема 26 кадра переключается на прием данных широковещательного канала.

В сетях стандарта 3G-4G закономерным развитием широковещательной передачи при ЧС являются не только системы PWS, но и системы, основанные на передаче сообщений через Интернет, поскольку нет необходимости в подключении и сотрудничестве с операторами сотовой связи.

### **Библиографический список**

1. **J. Eberspächer, H.-J . Vögel, C. Bettstetter, C. Hartmann.** GSM – Architecture, Protocols and Services Third Edition.: John Wiley & Sons, 2009.
2. **Fabian van den Broek.** Catching and Understanding GSM-Signals.: Radboud University Nijmegen, 2010.
3. **Ю.М.Туляков, М.В.Ерютов, А.Е. Шахов.** Анализ вариантов реализации широковещательной передачи данных (сообщений) в сотовой радиосвязи от 2G до 5G, Материалы XXVI международной научно-технической конференции «Информационные системы и технологии» (ИСТ-2020), Н.Новгород: НГТУ, 2020, С.178-182
4. [https://osmocom.org/projects/cellular-infrastructure/wiki/Cell\\_Broadcast](https://osmocom.org/projects/cellular-infrastructure/wiki/Cell_Broadcast)
5. <https://osmocom.org/projects/baseband/wiki/Phones>
6. [https://osmocom.org/projects/osmotrx/wiki/LimeSDR\\_Family](https://osmocom.org/projects/osmotrx/wiki/LimeSDR_Family)
7. <https://blog.3g4g.co.uk/search/label/PWS>
8. <https://habr.com/ru/news/t/415013/>
9. [https://www.rbc.ru/technology\\_and\\_media/14/09/2021/61408f739a79473e911b04fa](https://www.rbc.ru/technology_and_media/14/09/2021/61408f739a79473e911b04fa)

## ТЕХНОЛОГИЯ MMWAVE В СЕТЯХ ПЯТОГО ПОКОЛЕНИЯ. ВОЗМОЖНОСТИ И ПЕРСПЕКТИВЫ.

А.В. Пронин

ORCID: 0000-0002-9692-9556 e-mail: andreytemnikov.pronin@yandex.ru

Нижегородский государственный технический университет им. Р.Е. Алексеева

В докладе рассматривается технология передачи данных mmWave в сетях пятого поколения. Приведены ее общие принципы работы и характеристики, преимущества и проблемы развертывания в условиях городской среды. Проведен анализ и выявлены проблемы внедрения построения сетей на примере строений различной конфигурации. Описаны различные варианты построения сетей с данной технологией и сделаны соответствующие выводы. Помимо этого, было проведено исследование на предмет того, что сможет ли mmWave использоваться для решения проблемы последней мили.

**Ключевые слова:** mmWave, 5G, скорость передачи, NewRadio.

## MMWAVE TECHNOLOGY IN 5 GENERATION NETWORKS.

A.V. Pronin

Nizhny Novgorod state technical university n.a. R.E. Alekseev is a future-oriented university

**Purpose:** The study of mmWave technology, its prospects, methods of implementation and deployment problems in a city environment.

**Design / methodology / approach:** Metadology consists in collecting up-to-date information about the problem and its analysis, followed by a detailed analysis and possible proposals for its improvement. Using the developments and information of authors from countries where mmWave technology is being actively implemented or similar commercial networks are already operating, it is possible to identify certain patterns that will undoubtedly be useful when introducing technology in countries where it is just being developed. This will serve as a good basis for avoiding large outlays and failures.

**Findings:** In the work done, the results of the analysis of possible problems that will be encountered when deploying mmWave technology in various buildings and beyond are presented.

**Research limitations/implications:** Insufficiency of information in this area, or rather, numerical values with direct experimentation and their number. There is little information about the operation of networks with this technology in buildings of various configurations and from various materials to obtain an array of statistical data.

**Originality/value:** mmWave technology in 5G networks is currently in its infancy, despite the fact that it appeared along with new generation networks. Few countries can boast fully operational commercial networks with this technology, so the information provided in this paper certainly has high levels of value and originality. The area in which this study is being conducted contains a lot of unsolved problems.

**Key words:** mmWave, 5G, transmission speed, NewRadio.

## Введение

Технология mmWave позволяет достичь скорости передачи данных порядков в десятки Гбит/с такие значения даже в теории недостижимы для LTE. Использование высоких частот более 24 ГГц позволяет организовать сверхвысокоскоростные каналы связи достигая задержек при передаче сигнала порядка 1 мс, они оптимально подходят для предоставления инновационных услуг, которые развиваются в настоящее время. Однако за столь привлекательной скоростью и задержкой скрываются масса недостатков. Например, небольшая зона покрытия и существенные затухания сигнала при прохождении препятствий, учитывая тот факт, что рука человека уже представляет серьезное препятствие для миллиметровых волн поиск оптимальных решений для применения технологии mmWave стоит очень остро. На практике из-за небольшой зоны покрытия экономически целесообразно располагать БС в местах с наибольшим скоплением абонентов (UE устройств). Поэтому логично располагать БС в таких местах как: стадионы, торговые центры, спорткомплексы, офисы, образовательные учреждения, аэропорты и другие социально и культурно значимые места. В своей массе все эти объекты довольно шаблонны с инженерной точки зрения и отличаются лишь архитектурным решением. Как пример стадионы или театры, они являются по сути местностью с прямой видимостью и не имеют серьезных препятствий для прохождения сигнала, таких как перегородки или стены из бетона, дерева или стекла. Все зрители находятся в прямой видимости и поэтому применение mmWave при таких конфигурациях зданий ничем не усложнено, а расчет мест установки БС в основном будет сводиться к расчету зон покрытия сети. Однако с другими типами построек, как например с офисами все гораздо сложнее и требуется анализ как пути прохождения сигнала, так и возможных препятствий. Извилистые коридоры и разбитие пространства этажа на комнаты не позволят беспрепятственно пересекать их сигналу в миллиметровом диапазоне.

## Основная часть

Сеть нового поколения 5G в mmWave использует следующие диапазоны частот n258 (24.25 – 27.50 GHz), n257 (26.50 – 29.50 GHz), n261 (27.50 – 28.35 GHz), n260 (37.00 – 40.00 GHz), n259 (39.50 – 43.50 GHz), n262 (47.20 – 48.20 GHz) [1], распространение волны в данных диапазонах имеет свою специфику т.к. в существенную роль начинает играть физический процесс поляризации диэлектриков при прохождении через них сигналов, поэтому появляются новые параметры при расчетах такие как диэлектрическая проницаемость и тангенс угла диэлектрических потерь. Диэлектрическая проницаемость ( $\epsilon$ ) [2] это мера электрической поляризуемости диэлектрика. Например, материал с высокой диэлектрической проницаемостью поляризуется больше в ответ на приложенное электрическое поле, чем материал с низкой диэлектрической проницаемостью, тем самым накапливая больше энергии в материале. Само определение показывает, что чем больше диэлектрическая проницаемость материала, тем больше энергии он запасет проходящей через него волны, как следствие возрастают затухания. Тангенс угла диэлектрических потерь ( $tg\delta$ ) характеризует потери энергии электрического поля, которые рассеиваются в диэлектрическом материале. Обе величины дают представление о потерях волновой энергии, в табл. 1 приведены параметры материалов, которые применяются при строительстве зданий и отделки внутренних помещений.

Таблица 1.

Материалы, часто применяемые в ходе строительных работ и их диэлектрическая проницаемость

Наименование	Диэлектрическая проницаемость ( $\epsilon$ )
Бетон	4,5
Дерево	2,04-7,3
Стекло	3,8-19
Кирпич	4
Полипропилен	2,3-2,9
Резина	7,0
Воздух	1

Указанные в таблице параметры заметно разнятся, что явно говорит, а различном уровне затуханий проходящего через них полезного сигнала[2]. В рамках данной работы предполагается выявить наиболее перспективные методы построения сети связи внутри здания с использованием технологии mmWave. При этом ставится задача использовать уже доступные и рабочие технологии.

В первыйнаиболее простой и понятный метод можно выделитьиспользование в качестве межкомнатных разделителей радиопрозрачных материаловдля лучшего прохождения сигнала, тем самым снижаяего затухание и увеличивая зону покрытия. Радиопрозрачные материалы – это диэлектрики, которые не изменяют амплитуду электромагнитной волны, проходящей через них, и не вызывают хаотичного изменения ее фазы. Прозрачность этих материалов для радиоволн означает, что в интервале рабочих температур они имеют очень малые диэлектрические потери и практически не отражают радиоволн.В основном подобные материалы используют для защиты различных антенн от пагубного влияния внешних сил в военной техники или довольно специфичных задачах, поэтому на данный момент и в ближайшем будущем не предвидится массового производства подобных материалов, следовательно, и применение в строительстве в качестве отделочных материалов в ближайшие годы маловероятно. Для оценки затухания волны при прохождении через препятствие можно воспользоваться первым членом решения уравнения Гельгольца:

$$\sigma = e^{-k_{\alpha} l}$$

где  $l$ -расстояние пройденное волной, в нашем случае это толщина материала, а  $k_{\alpha}$ - волновое число зависящее от диэлектрической и магнитной проницаемости, его можно посчитать используязаписи комплексных проницаемостей среды:

$$\tilde{\epsilon}_{\alpha} = \epsilon'_{\alpha} - i\epsilon''_{\alpha} = \epsilon'_{\alpha}(1 - itg\delta^{\epsilon})$$

$$\tilde{\mu}_{\alpha} = \mu'_{\alpha} - i\mu''_{\alpha} = \mu'_{\alpha}(1 - itg\delta^{\mu})$$

где  $tg\delta^{\epsilon} = \frac{\epsilon''_{\alpha}}{\epsilon'_{\alpha}}$ ,  $tg\delta^{\mu} = \frac{\mu''_{\alpha}}{\mu'_{\alpha}}$ , тогда:

$$\tilde{k} = \omega\sqrt{\tilde{\epsilon}_{\alpha}\tilde{\mu}_{\alpha}} = k_0\sqrt{\tilde{\epsilon}} = k_0\sqrt{\epsilon - i60\lambda_0\sigma} \quad (1)$$

$$k_{\beta} = k = \omega\sqrt{\epsilon_{\alpha}\mu_{\alpha}} \quad (2)$$

С учетом (1) и (2)  $k_{\alpha}$  находится как:

$$k_{\alpha} = \frac{k}{2}tg\delta = \frac{\omega}{2}\sqrt{\epsilon_{\alpha}\mu_{\alpha}}tg\delta$$

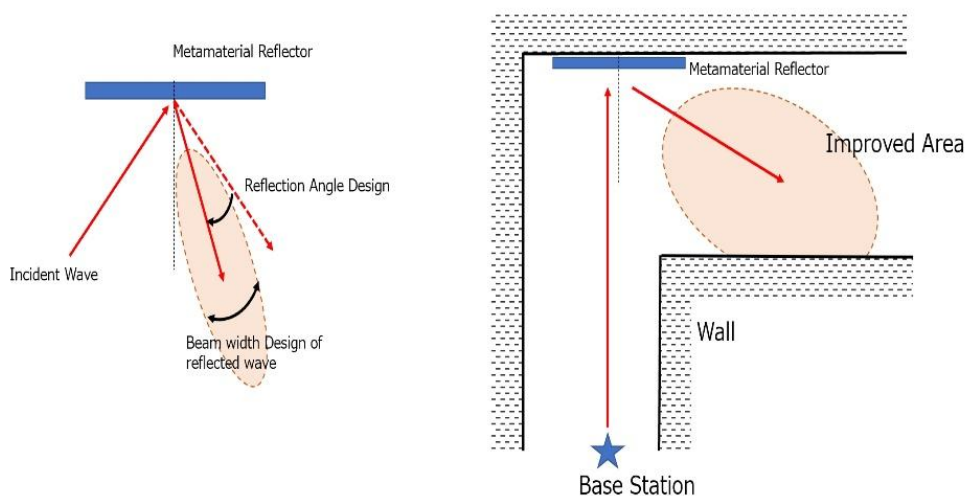
Используя произведение  $\prod_1^n \sigma_n$ , где  $n$  число препятствий,можновычислить затухание сигнала и определить область покрытия задавая минимально допустимый уровень.

Второйметод прямо противоположный первому - проектирование помещенийиспользуя радиопоглощающие материалы в комбинации с отражателями. Основная особенность радиопоглощающих материалов заключается в их способности преобразовывать электромагнитную волну внутри их структуры в иные виды энергии. При этом происходят явления, такие как: поглощение и рассеяние, интерференция и дифракция.Существуют различные виды материалов, которые могут поглощать энергию волны, и они пригодны для отделки внутри помещений, например, различные роды пластмасс и каучука, многослойная керамика и стекло с ферромагнитными примесями. Метод предполагает использовать поглощение и отражение сигнала от поверхностей, используя технологию формирования луча (Beamforming)при проектировании этажа здания необходимо предусмотреть специальные проёмы между отдельными помещениями, материал которых будут иметь хороший коэффициент отражения. Технология формирования узконаправленного луча применяется для увеличения эффективности использования спектра и количества каналов, она позволяет сформировать узкий луч с массива антенн тем самым направить передачу непосредственно к принимающему устройству. Учитывая, что миллиметровые радиоволны обладают хорошими показателями направленности за счет

увеличения в разы количества антенных элементов на антенну mmWave дает огромный выигрыш этой технологии:

- мощный сигнал на выходе в направлении к UE;
- сильный уровень сигнал/шум в направлении от UE;
- отсутствие межсотовой интерференции;
- значительное увеличение количества каналов связи на одну соту.

С помощью специальных вычислителей меняя угол направления луча можно добиться отражений и переотражений от отражателей таким образом, чтобы достичь нужной точки в здании, что продемонстрировано на рис.1. На первый взгляд выглядит фантастично и не реализуемо, однако еще в 2017 году на Петербургском экономическом форуме компания Мегафон провела эксперимент используя широкую полосу частот от 2 ГГц до 70 ГГц, ей удалось достичь скорости передачи данных в 35 Гб/сек [3]. В эксперименте Мегафона был один излучатель и сложной формы антенна, которая формировала несколько пучков — два параллельных и один под углом, специальная пластина из отражающего материалы была расположена на полу. Опыт увенчался успехом исходя из этого управление отраженными сигналами можно считать вполне реализуемыми. Метод с использованием узконаправленного луча к сожалению, имеет один нежелательный эффект в виде боковых лепестков в диаграмме направленности во время формирования многолучевости при отражениях, однако и эти проблемы решаются за счет использования поглощающих материалов, которые в разы или практически полностью поглощают образовавшиеся помехи [4-7]. Казалось бы, данный метод имеет массу преимуществ, но у различных радиопоглощающих материалов степень поглощения энергии разная и зависит от конкретной частоты волны. Поэтому не получится использовать однотипные материалы для поглощения волн в широком диапазоне частот, а придется использовать их комбинации при проектировании помещения с учетом конкретных диапазонов частот.



**Рис. 1. Отражение сформированного луча от рефлектора**

Для контроля каналов при построении сети по такому принципу можно воспользоваться технологией SRS (Sounding Reference Signal – излучение опорного сигнала). Она используется дополнение к формированию луча, позволяющее базовой станции узнать о качестве канала через специальный пакет, посылаемый от UE. На практике большинство UE могут поддерживать отправку SRS только через свою основную передающую антенну таким образом БС может получать информацию только о канале для этой антенны. Однако, при использовании технологии выбора передающей антенны, можно получить полную информацию о каналах всех антенн UE. Благодаря этому, БС получит возможность генерировать луч в направлении UE наилучшим образом. В результате прирост пропускной способности для UE значительно увеличится, особенно в точках на дальнем и среднем удалении от базовой станции [8].

Третьей методобеспечения без разрывного покрытия сети в миллиметровом диапазоне внутри помещения может быть размещение малых сот по всему зданию. Это возможно благодаря недавно представленному решению для сотовой связи 5G внутри помещений - Link Cell

(Samsung). Устройство поможет сделать связь бесшовной и расширить возможности 5G для пользователей внутри помещений. По задумке авторов использоваться оно будет на предприятиях, включая производственные или распределительные предприятия, корпоративные офисы, а также развлекательные или общественные места. Link Cell является одним из первых коммерческих продуктов, доступных во всем мире, который позволяет операторам беспроводной связи организовать малые соты для помещений миллиметрового диапазона, на рис. 2 показано его реальное изображение [9]. Первая версия устройства будет поддерживать частоту 28 ГГц и имеет возможность комбинировать четыре полосы пропускания частот по 100 МГц, предлагая высокую емкость и сверхбыструю скорость загрузки. Кроме того, в его составе уже включены все необходимые схемотехнические компоненты как приёмопередающие антенны, блоки интерфейсов и цифровой обработки, при этом все выполнено в миниатюрных габаритах, что обеспечивает быструю и простую установку внутри помещений, например, на стене или потолке.



**Рис. 1. Общий вид базовой станции для построения малой соты с использованием mmWave**

Все БС малых сот с помощью транспортной сети, например, оптоволоконной соединяются в конечном едином распределительном устройстве (HUB) с которого и происходит обмен информацией с каждой БС в здании. В этом моменте появляется ключевой недостаток учитывая, что для покрытия всех активных зон необходимо большое количество малых сот, даже с учетом алгоритмов распределения доступа понадобится внушительное количество проводных линий. Поэтому в затратах на проектирование по мимо дорогих БС необходимо будет учитывать и обширную транспортную сеть [10-12].

Как известно, самым узким местом во всей структуре сетей передачи информации является участок сети от абонента до ближайшего узла коммутации - «последняя миля». Полоса пропускания на этом участке по сравнению с магистральными сетями очень мала, что ограничивает возможности по предоставления абоненту многих современных услуг таких, как высокоскоростной доступ в Интернет, IPTV высокой четкости, AR/VR и т.д. Существует множество решений этой проблемы, но большинство из них довольно затратно потому, что связано с наращивание новых абонентских линий. Хотя технология FTTx практически решает эту проблему для абонентов неподвижной связи, но для мобильных устройств вопрос все еще актуален и как никогда становится острым, с учетом роста потребности в мобильные передачи информации диапазон миллиметровых волн вполне может решить эту проблему. Сети, построенные с технологией mmWave, имеют гораздо большую емкость и скорости передачи. Например, для того чтобы воспроизводить потоковое видео в быстро набирающем популярность формате 2160p (4K) понадобится скорость около 85 Мбит/с (зависит от битрейта, HDR и д.р.). Как говорилось выше данный диапазон обеспечивает скорости в десятки Гбит/с, они с легкостью покроют все потребности пользователей.

### **Выводы**

Были предложены три потенциально вероятных метода для построения сети связи внутри здания с использованием миллиметровых длин волн. Каждый из методов был предложен с учетом

текущего уровня доступных технологий для его реализации. Рассмотрены их преимущества, недостатки, а также перспективы развития и реализации на практике.

Первый метод, который предполагает использование радиопрозрачных материалов является самым простым в реализации в теории, однако на практике из-за неразвитости производства подобных материалов является самым затруднительным. На данный момент производство таких материалов обходится дорого, все потому что это довольно специфичная область в которой выполняются единичные проекты. Несмотря на это судя по заявлениям инженеров в области телекоммуникации во всем мире, о том, что недавно вышедшая связь нового поколения 5G, а именно ее диапазон частот ниже 6ГГц уже полностью распределён и в ближайшие годы будет освоен, а также о том, что необходимо переходить в область частот выше 24ГГц, дают четко понять о возможном наращивании темпов производства материалов для работы со столь высокими частотами для повсеместного использования.

Второй метод гораздо сложнее осуществим в теории и на практике по отношению к первому. В нем задействованы комбинации сразу несколько технологий, которые по одиночке не представляют из себя проблем при использовании. Однако их совместное применение не тривиальная инженерная задача, к тому же такой подход потребует хороших дополнительных вычислительных мощностей по мимо мощности для обслуживания сети связи. Хотя приведенные недостатки и кажутся существенными, большая зона обслуживания вместе с высоким SINR за счет использования сформированного узконаправленного луча и эффекта поглощения, делают этот метод достойным рассмотрения. Еще одним существенным минусом может оказаться то, что не получится обойтись простыми расчетами, а необходимо проводить полноценное моделирование. Помимо того, что понадобится дополнительная рабочая сила из-за необходимости детального переноса исследуемого помещения в электронную модель, так еще и популярные программные продукты для расчетов распространяются по очень дорогой временной лицензии.

Третий метод использование готовых решений, а именно множества БС для построения малых и возможно сверхмалых сот можно отнести к фаворитам. Готовые решения имеют малые габариты и все необходимое для развертывания сети с использованием mmWave в здании. По мимо этого они имеют интерфейсы для ручной и автоматизированной настройки, а также вторичные радио интерфейсы, позволяющие организовывать служебные каналы связи с UE устройствами. Существенным недостатком этого метода являются несоизмеримые финансовые затраты, как на сами БС, так и на обширную транспортную сеть, дополнительное вычислительное и распределительное оборудование.

Без сомнения, сети связи mmWave могут стать решением проблемы последней мили. Скорость передачи в десятки Гбит/с и малые задержки позволят абонентам получать самые современные услуги, которые может им предложить поставщики.

#### Библиографический список

1. 5G Frequency Bands // HBR URL: <https://halberdbastion.com/technology/cellular/5g-nr/5g-frequency-bands> (дата обращения: 10.05.2022).
2. Диэлектрическая проницаемость // энергетик URL: <https://energetik.com.ru/zakony-elektrotexniki-korotko/dielektricheskaya-pronicaemost> (дата обращения: 10.05.2022).
3. 35 Гбит/с — как Megafon и Huawei поставили рекорд скорости 5G // Хабр URL: <https://habr.com/ru/company/megafon/blog/337204/> (дата обращения: 10.05.2022).
4. Диэлектрическая проницаемость различных веществ, в т.ч. основных диэлектриков. // Технические таблицы URL: <https://tehtab.ru/Guide/GuidePhysics/ElectricityAndMagnethism/DEPLiquids/DielectricConstanVarious/> (дата обращения: 10.05.2022).
5. IMPRESS: Indoor Mobility Prediction Framework for Pre-Emptive Indoor-Outdoor Handover for mmWave Networks // IEEE Xplore URL: <https://ieeexplore.ieee.org/document/9642053> (дата обращения: 10.05.2022).
6. Millimeter-Wave Propagation Measurement and Modeling in Indoor Corridor and Stairwell at 26 and 38 GHz // IEEE Xplore URL: <https://ieeexplore.ieee.org/document/9435360> (дата обращения: 10.05.2022).

7. Coverage Enhancement for NLOS mmWave Links Using Passive Reflectors // IEEE Xplore URL: <https://ieeexplore.ieee.org/document/8972365> (датаобращения: 10.05.2022).
8. Indoor Office Wideband Millimeter-Wave Propagation Measurements and Channel Models at 28 and 73 GHz for Ultra-Dense 5G Wireless Networks // IEEE Xplore URL: <https://ieeexplore.ieee.org/document/7289335> (датаобращения: 10.05.2022).
9. Samsung Brings 5G Indoors with New Commercial 5G mmWave Small Cell for In-building Use // Samsung Newsroom U.S. URL: <https://news.samsung.com/us/5g-indoor-mmwave-small-cell/> (датаобращения: 10.05.2022).
10. Millimeter Wave Networked Wearables in Dense Indoor Environments // IEEE Xplore URL: <https://ieeexplore.ieee.org/document/7445132> (датаобращения: 10.05.2022).
11. Fingerprinting-Based Indoor Localization With Commercial MMWaveWiFi: A Deep Learning Approach // IEEE Xplore URL: <https://ieeexplore.ieee.org/document/9081910> (датаобращения: 10.05.2022).
12. Comparative Study of Artificial Neural Network Based Channel Equalization Methods for mmWave Communications // IEEE Xplore URL: <https://ieeexplore.ieee.org/document/9374910> (датаобращения: 10.05.2022).



**РАЗРАБОТКА ПРОГРАММЫ АНАЛИЗА  
УРОВНЯ ЗНАНИЙ ОБУЧАЮЩИХСЯ**

Д.А.Шевченко

ORCID: 0000-0000-0000-0000 e-mail: d.shevchenko1@yandex.ru

Нижегородский государственный технический университет им. П.Е. Алексеева

В работе предложена разработка одной из разновидностей систем анализа уровня знаний обучающихся, реализованная с помощью ИРТ систем, сетевых систем, а также модели Раша, которая позволяет преобразовать полученные данные тестирования в некоторые оценки. Система позволяет оценивать измерение уровня подготовки пользователя системы во времени. Используется линейная шкала параметров уровня знаний. Для обеспечения достоверности оценок испытуемых нужно отслеживать качество самих тестовых заданий. Основной задачей ИРТ в данной программе является переход от индикаторных переменных к латентным параметрам. В ИРТ устанавливается связь между двумя множествами значений латентных параметров. Метод анализа тестовых задач с помощью модели Раша дает обширные, объективные и достаточно полные знания о обучающихся, выдает статистически обоснованные результаты анализа, а также позволяет прогнозировать детальную диагностику. Инструменты реализации GUI предоставляют мощный массив элементов управления GUI. Для настройки программного обеспечения дизайнеры могут изменить код соответствующим образом.

**Ключевые слова:** ИРТ, модель Раша, GUI, сетевые модели.

**DEVELOPMENT OF THE ANALYSIS PROGRAM  
LEVEL OF KNOWLEDGE OF STUDENTS**

D.A.Shevchenko

Nizhny Novgorod state technical university. R.E. Alekseeva

**Purpose:** development of one of the varieties of systems for analyzing the level of knowledge of students, implemented using IRT systems, network systems, as well as the Rasch model.

**Design / methodology / approach:** The Rasch model assumes a probabilistic dependence on susceptibility (to the subject) according to the manifested inclination to the indicator on the level of knowledge in general, i.e. the more pronounced in the object the measured latent quality as a whole, the more it should be expressed in all indicators. The main postulate of the model: a subject with a more pronounced increase in the level of consciousness shows greater activity in all respects than a subject with a less pronounced degree. The fulfillment of this condition is a consequence of the emergence of one or another indicator of it and corresponds to the logical basis of the model. As a consequence of what was predicted above, the indicator should have differentiating sensitivity, i.e. e. to divide the subjects according to the level of knowledge, but to varying degrees.

**Findings:** The method of analyzing test problems using Rasch models provides extensive, objective, and fairly complete knowledge of the fundamental ones, provides statistically valid information about the analysis, and also makes it possible to predict a detailed problem.

**Research limitations/implications:** The study is limited to a database that includes test items, answers, and a theoretical base.

**Originality/value:** The task of determining the level of knowledge of trainees, as well as analyzing the correctness of practical tasks, is extremely relevant at the present time. In the educational process, it is necessary to control the quality of the taught matter, and with it an assessment, to determine the level of knowledge.

**Key words:** IRT, Rasch model, GUI, network models.

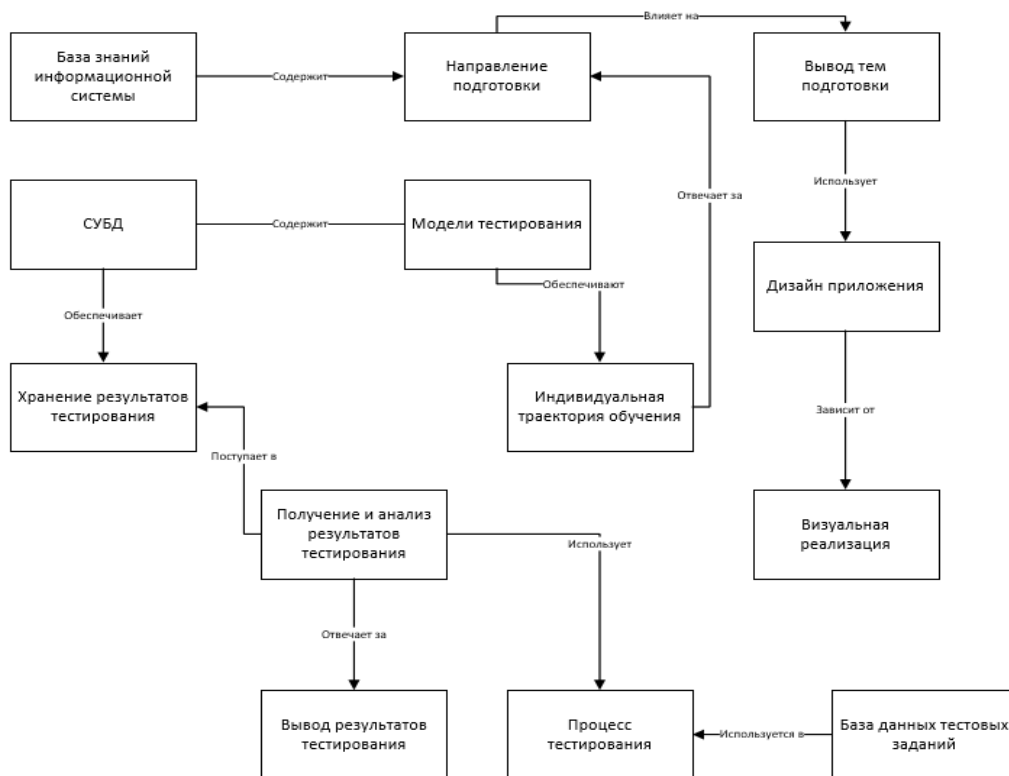
В настоящее время тяжело представить процесс обучения в образовательных учреждениях без использования обучающих компьютерных программ и технологий тестирования усвоенной информации. Каждый год все больше обучающих учреждений внедряют системы тестирования, улучшают образовательный процесс и ускоряют процесс проверки знаний. Также все больше преподавателей стали использовать компьютер как средство подготовки и оценки знаний своих учеников.

В данный момент существует большое количество программ для тестирования, позволяющих не только создать и провести компьютерное тестирование, но также и после прохождения теста получить отчет об уровне усвоения информации учащимися, а также сам ученик имеет доступ к своим тестовым работам. Основной задачей подобных систем является доступность информации, для обучения студентов, а также удобство в проверке полученных знаний.

При разработке базы данных, помимо определения типов связей между объектами предметной области, а в моделировании между сущностями, фиксируются дополнительные характеристики каждой связи. К таким характеристикам относятся мощность (кардинальность) связи, смысловое наполнение связи, правила ссылочной целостности.

Под мощностью связи понимается возможное количество экземпляров соответствующего объекта, связанных с одним экземпляром другого объекта (Рис.1).

Для определения связи необходимо проанализировать базы данных и вспомогательные сведения из предметной области, взаимосвязи двух рассматриваемых структурных элементов. Этот анализ должен показать количество значений одного информационного объекта в связи с одним экземпляром значения другого объекта. При этом необходимо рассматривать связь в обоих направлениях: от первого информационного объекта ко второму и наоборот. Подобный анализ связи позволит определить верхнюю границу мощности, но не даст практически никаких сведений о нижней границе.



**Рис 1. Блок-схема проектирования системы**

## Модель Раша

Тест, как средство измерения учебных достижений, может дать достоверный результат только в случае его корректного применения. Корректность применения теста – это многоаспектное понятие, включающее в себя вопросы конструирования и дизайна теста, вопросы разработки и применения тестов и, разумеется, интерпретации результатов тестирования. В данной работе основное внимание уделено вопросам корректности интерпретации результатов педагогического тестирования, проводимого на основе модели Г.Раша. Анализ результатов обычно проводится на основе классической теории тестов или на основе Item Response Theory..

Достоинства:

- Создана единая шкала интервалов, для измерения уровня подготовки, а также уровня знаний.
- Отсутствует зависимость теста и оценки уровня подготовки
- Присутствует возможность увидеть соотношение любого тестируемого с заданием, с помощью чего можно будет предположить вероятность удачной сдачи теста данным испытуемым.

Так, благодаря построенной системе был сделан вывод, что данная разработанная программа позволяет составителям тестов качественно и быстро оценить инструментарий по трудности, выявить задания, не соответствующие принятым нормам и требованиям, а так же в ходе тестирования получить объективные оценки испытуемых.

Контроль качества тестовых заданий является неотъемлемой частью процесса улучшения работы образовательных учреждений. Есть много различных способов создания систем оценки тестирований, например такие как : INDIGO, I-exam, АСТ-тест и др.

Сам же математический аппарат однопараметрической модели Раша будет приведен ниже. Были обозначены уровни подготовки испытуемых как  $\theta_i$  ( $i=1 \dots n$ ), где  $n$  – количество испытуемых, а сложность заданий теста –  $\beta_j$  ( $j=1 \dots m$ ), где  $m$  – количество вопросов. Теория Раша устанавливает связь между множествами параметров  $\theta_i$  и  $\beta_j$ .

Причем  $\theta_i$  и  $\beta_j$  оцениваются в одной и той же шкале, а функция успеха испытуемого имеет в качестве аргумента разность  $(\theta_i - \beta_j)$ . Если  $(\theta_i - \beta_j) > 0$  и велика, вероятность успеха  $i$ -го испытуемого в  $j$ -м задании велика. При условии, что  $(\theta_i - \beta_j) < 0$ , но  $|\theta_i - \beta_j|$  велик, вероятность успеха  $i$ -го испытуемого в  $j$ -м задании мала. В качестве единой шкалы Раш предложил ввести интервальную шкалу логитов.

Интервальной является шкала, в которой значимыми считаются расстояния между измеряемыми величинами. Логитом названа единица измерения уровней готовности участников тестирования и сложности тестовых заданий в рамках логистических моделей тестирования (от *logo* – слова, речь). В общем виде преобразование шкалы логитов имеет вид:

$\theta_1 = \alpha + \gamma\theta$  (2), где  $\theta_1$  – оценки уровней подготовки тестируемых,  $\theta$  – оценки параметров в шкале логитов,  $\alpha$  – константа, служащая для определения начала шкалы (константа переноса),  $\gamma$  – коэффициент шкалирования для определения размерности шкалы.

$\beta_2 = \alpha + \gamma\beta$ , где  $\beta_2$  – оценки уровней трудности заданий,  $\beta$  – оценки параметров в шкале логитов,  $\alpha$  – константа переноса,  $\gamma$  – коэффициент шкалирования.

Преобразование исходного балла направлено на то, чтобы убрать из оценки тестируемого зависимость от трудности заданий теста и из трудности заданий зависимость от конкретной выборки испытуемых.

Модель Раша позволяет преобразовать все полученные данные в некоторые оценки, которые можно рассматривать как результат измерения уровня подготовки пользователя системы. Используется линейная шкала параметров уровня знаний. Для отслеживания достаточно достоверных оценок испытуемых нужно отслеживать качество самих тестовых заданий (рис.2).

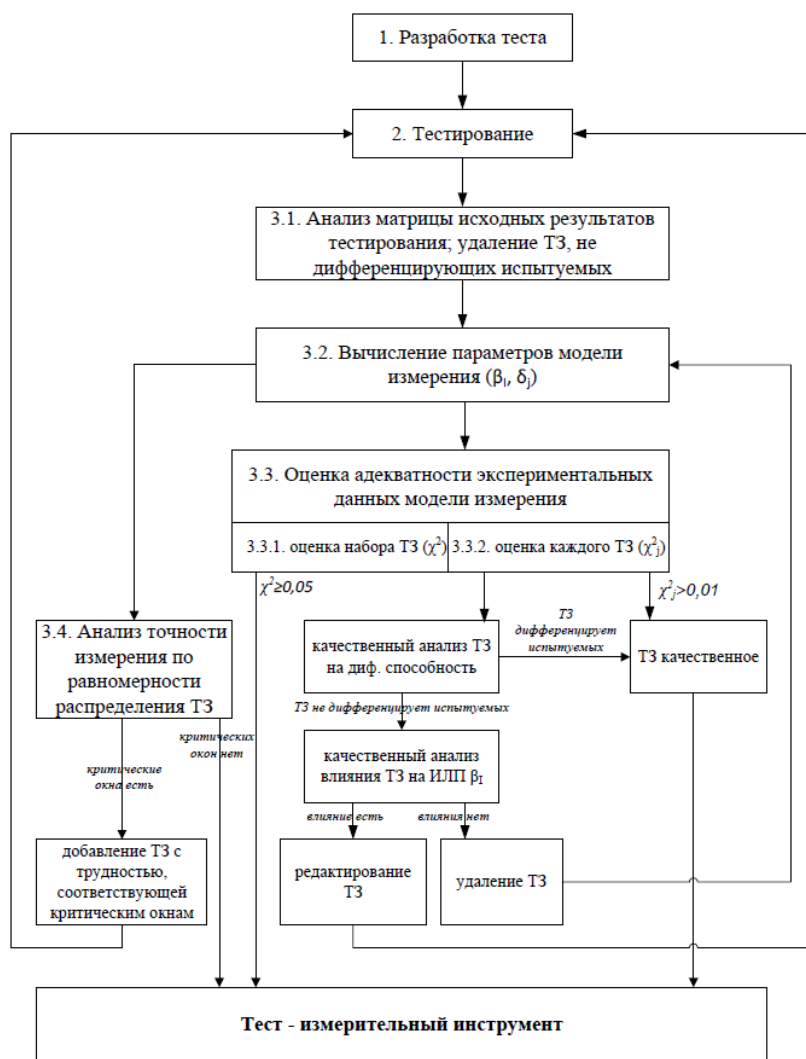


Рис 2. Алгоритм проверки результата тестовых заданий

Из наиболее популярных программ для тестирования можно отметить:

1. MyTestXPro.

Комплекс программ (программа тестирования учащихся, редактор тестов и журнал результатов) для проведения компьютерного тестирования, сбора и анализа результатов, выставления оценки по указанной в тесте шкале.

2. NetTest.

Программный комплекс NetTest предназначен для проведения удаленного тестирования знаний. Комплекс состоит из двух программ – серверной и клиентской. Серверная программа позволяет создавать и редактировать тесты, она также осуществляет полное управление компьютерным тестированием, обработку и вывод результатов. Клиентская программа запускается с рабочих станций и предназначена для работы учащегося.

3. «АСТ-Тест».

Комплекс программ для создания банков тестовых заданий, организации и проведения тестирования, обработки результатов и формирования выходных документов в удобной форме.

4. «Айрен».

Это бесплатный программный комплекс, позволяющий создавать тесты для проверки знаний и проводить тестирование в локальной сети, через Интернет или на одиночных компьютерах.

Программные комплексы включают в себя механизмы тестирования различных типов: с выбором одного или нескольких верных ответов, с вводом ответа с клавиатуры, на установление соответствия, на упорядочение и на классификацию, а также некоторые программные комплексы поддерживают аудио или видео-задания. Сравнение данных систем представлено в табл. 1.

Сравнение существующих систем тестирования

Критерий	MyTestXPro	NetTest	«АСТ - тест»	«Айрен»
Вопросы одним верным ответом	+	+	+	+
Вопросы с несколькими верными ответами	+	–	+	+
Вопросы с вводом ответа	+	+	+	+
Вопросы на установление соответствия	+	–	+	+
Вопросы с аудио/видео - содержанием	+	–	+	–

При рассмотрении данных программных комплексов проявились такие недостатки, как недостаточное или полное отсутствие интерактивности в тестировании, дизайн приложений выглядит довольно скупо, практически отсутствует работа с графикой.

Модель бизнес-процесса есть описание порядка выполнения работ, приводящих к достижению вполне определенного и воспроизводимого результата. Под бизнес моделью принято понимать формализованное (графическое, табличное, текстовое, символическое) описание бизнес-процессов. Модель должна давать полное, точное и адекватное описание системы. Модель бизнес-процесса – это абстрактное представление процесса, удовлетворяющая целям моделирования. В модели отображаются соответствующие (релевантные) аспекты бизнес-процесса, например, события, действия, порядок их выполнения. BPMN (Business Process Modeling Notation) – это графическая нотация для моделирования бизнес процессов (рис.3).

Стандарт BPMN является доступным в интернет документом. Он детально описывает особенности реализации нотации. Операция – это единица работы, выполняемая непрерывно, на одном рабочем месте, над одним обрабатываемым объектом. Действие есть акт взаимодействия оператора с обрабатываемым изделием, в котором достигается определенная, заранее определенная, цель. Выполнение операции приводит к качественным изменениям обрабатываемого изделия, а действия – к количественным.

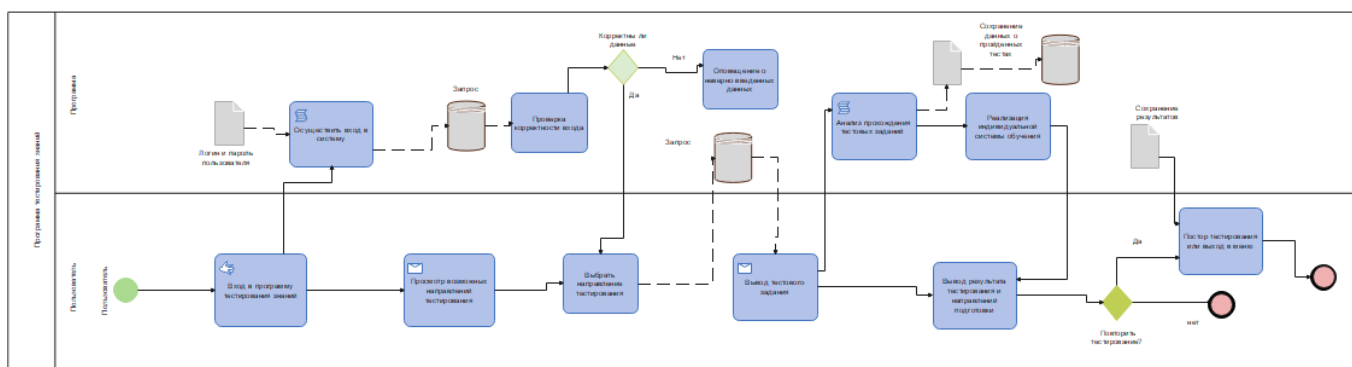
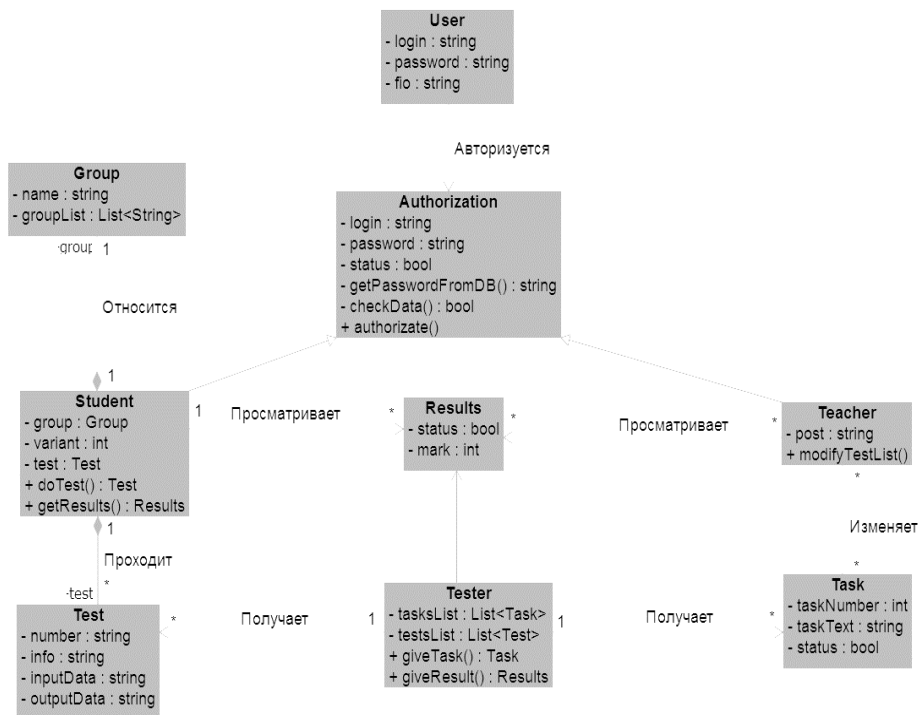


Рис.3 Схема BPMN

Типовой ход процесса работы системы: студент при начале работы программы тестирования знаний обращается к самому приложению. Этот запрос передается в СУБД, анализируется и просматривает уже существующие профили. После идет процесс входа в аккаунт и просмотр уже существующих сохраненных тестирований, если они имеются. Если тестирования уже были пройдены, то программа показывает пользователю какие темы были пройдены, результаты тестирований. Это осуществляется за счет обращения к БД с уже пройденными и сохраненными тестами, которые привязаны к аккаунту. После просмотра всей имеющейся информации пользователь может повторно выбрать тему, тестирование по которой он хочет пройти. Идет обращение к БД с тестовыми заданиями, выбирается нужная тема, а после этого запускается процесс тестирования. Его результаты передаются для анализа.



**Рис. 4. Диаграмма классов**

При выборе средства реализации среди множеств платформ я остановилась на веб-реализации интерфейса.

Плюсы реализации веб – приложения:

- С веб-приложением можно работать в любой точке мира с компьютера, планшета или смартфона, подключенного к Интернету.

- Экономия

Веб-приложения работают на всех платформах и исключают необходимость разработки приложения отдельно для Android и iOS.

- Адаптивность

Если для нативных приложений нужны определенные ОС, то для работы с веб-приложением подходит любая ой операционной системой (Windows, MAC, Linux и т.д.) и любой браузер (Internet Explorer, Opera, FireFox, Google Chrome и т.д.).

- Отсутствие клиентского ПО

Дешевле и проще установка, обслуживание и модернизация клиентского интерфейса. Обновление до последней версии происходит при очередной загрузке страницы.

- Сетевая безопасность

Веб-система имеет единую точку входа, защитить и настроить безопасность которой можно централизованно.

- Защита от потери данных

Данные пользователей хранятся в «облаке», за целостность которого отвечают хостинг-провайдеры, и защищены от потери при повреждении жесткого диска компьютера.

Студент имеет возможность поэтапно отправлять задания на проверку. Например, как только он закончил решать одно задание, то сразу может отправить его на проверку. Это дает преподавателю возможность проверять задание, не дожидаясь решения всей работы, что значительно сокращает время проверки всей работы.

При комплексной оценке показателей качества программного продукта качество пользовательского интерфейса определяется такой характеристикой качества, как юзабилити (usability).

Юзабилити (usability) – это степень, в которой продукт может быть использован определенными пользователями для достижения поставленных целей с учетом следующих критериев:

- эффективность (effectiveness) – точность и полнота, с которой пользователи достигают поставленных целей;
- экономичность (efficiency) – отношение израсходованных ресурсов, к точности и полноте, с которой пользователи достигают поставленных целей;
- удовлетворенность (satisfaction) – комфорт и приемлемость использования.

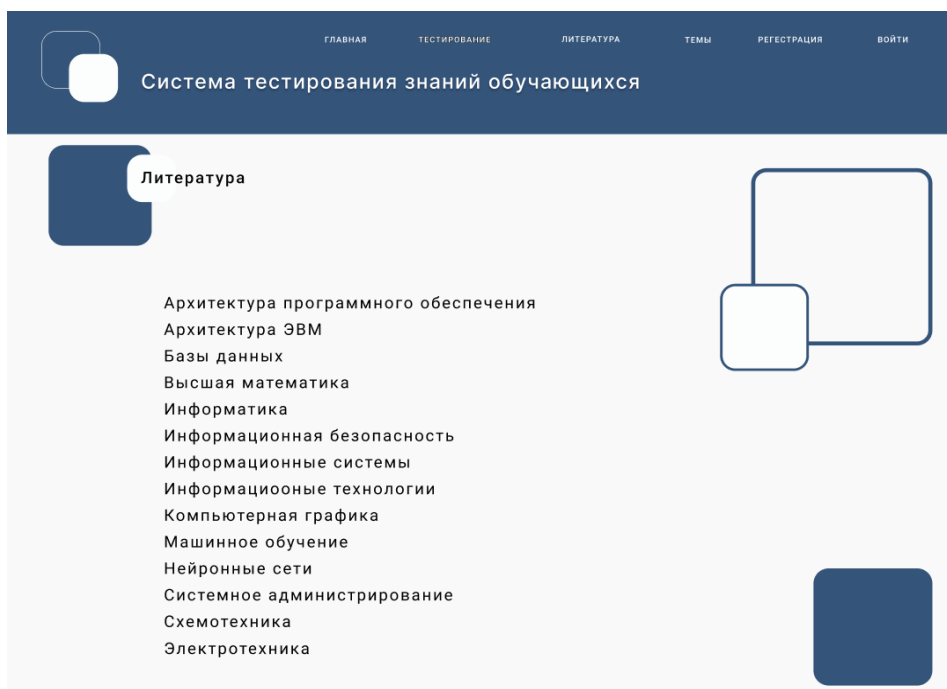
Разработка юзабилити – это методологический подход к созданию пользовательского интерфейса любой информационной системы. Это практический путь к получению продукта, который работает для пользователя.



**Рис. 5. Окно тестирования**

Студент не может редактировать решение задания после его отправки преподавателю, а преподаватель имеет возможность прокомментировать работу студента и при необходимости отправить ее на доработку, поставив статус «Ожидает исправлений».

Изменения в работе интерактивно отображаются у всех участников (студент-преподаватель). Так если преподаватель проверит отдельное задание и поставит оценку, то студент сразу увидит это у себя. Также у студента есть возможность просмотреть библиотеку нужной литературы в отдельной вкладке (рис. 5)



**Рис. 6. Экран со списком направлений электронной библиотеки**

## Выводы

Рассмотрев уже существующие системы тестирования и сделав выводы о их недостатках был создан алгоритм системы тестирования. Определены критерии оценки эффективности анализа системы тестирования, а также рассмотрены критерии оценки интерфейса системы. Сформированы базы данных, включающие в себя информацию о пользователях, списки литературы и тестовые задания.

## Библиографический список

1. **Хайдаров, КА.** 40-Технологии. Основы визуального программирования [Электронный ресурс]. URL: [http://bourabai.kz\]einf/4gl.htm](http://bourabai.kz]einf/4gl.htm) (дата обращения: 27.09.2016)
2. Что такое HiAsm? [Электронный ресурс]. URL: <http://www.hiasm.com/sdk.html> (дата обращения: 27.09.2016)
3. **Соловьев, С.В.** Технология разработки прикладного программного обеспечения / С. В. Соловьев, Р. И. Цой, Л. С. Гринкруг. – М. : Изд-во Академия Естественных наук, 2011. Режим доступа : <http://www.monographies.ru/141>.
4. **Гультяев, А. К.** Проектирование и дизайн пользовательского интерфейса / А. К. Гультяев, В. А. Машин. – СПб. : КОРОНА Принт, 2000. – 352 с.
5. ГОСТ Р ИСО 9241-210-2012 Эргономика взаимодействия человек – система. Часть 210. Человеко-ориентированное проектирование интерактивных систем. – М. : Стандартиформ, 2013. – 36 с.
6. ГОСТ Р ИСО/МЭК 9126-93 Информационная технология. Оценка программной продукции. Характеристики качества и руководства по их применению. – М. : Стандартиформ, 2004. – 12 с.



## ИНТЕЛЛЕКТУАЛЬНОЕ УПРАВЛЕНИЕ КАЧЕСТВОМ И ЭКОБЕЗОПАСНОСТЬЮ СОСТОЯНИЙ СЛОЖНЫХ ТЕХНИЧЕСКИХ И ТЕХНОЛОГИЧЕСКИХ ОБЪЕКТОВ

С.А. Манцеров

ORCID: 0000-0001-8458-8259 e-mail: mca\_9@nntu.ru

Нижегородский государственный технический университет им. Р.Е. Алексева

**Цель работы:** создание системы адаптивных моделей вычислений и технологий для интеллектуального управления качеством и экобезопасностью состояний сложных технических объектов и технологических процессов.

**Результаты:** система моделей вычислений, включающая «мягкие» и оптимизационные модели вычислений, адаптированная для решения задач обеспечения контролепригодности, отказоустойчивости и экобезопасности состояний сложных технических объектов и технологических процессов.

**Научный подход:** решение поставленных задач основано на применении теории системного анализа, обработки и систематизации информации, теории управления и методов принятия решений, теории нечётких множеств, методов математического и структурного моделирования, функциональной систематики.

**Новизна:** заключается в новой концепции интеллектуального управления качеством и экобезопасностью состояний технических и технологических объектов. В отличие от известных подходов к управлению качеством и экобезопасностью состояний технических и технологических объектов, неориентированных, как правило, на проблему ускорения вычислений, предложенная обобщенная схема интеллектуального управления качеством и безопасностью состояний технических и технологических объектов основана на модифицированных моделях вычислений и технологиях.

**Ключевые слова:** интеллектуальное управление, технические и технологические объекты, модели вычислений, экобезопасность, техническое состояние.

## INTELLIGENT QUALITY MANAGEMENT AND ENVIRONMENTAL SAFETY OF COMPLEX TECHNICAL AND TECHNOLOGICAL FACILITIES

S.A. Mantserov

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

**The purpose** of the work is to create a system of adaptive models of computing and technologies for intelligent quality management and environmental safety of complex technical objects and technological processes.

**Results:** a system of computing models, including "soft" and optimization models of computing, adapted to solve the problems of ensuring the controllability, fault tolerance and environmental safety of the states of complex technical objects and technological processes.

**Scientific approach:** the solution of the tasks is based on the application of the theory of system analysis, processing and systematization of information, management theory and decision-making methods, the theory of fuzzy sets, methods of mathematical and structural modeling, functional systematics.

**Novelty:** it consists in a new concept of intelligent quality management and environmental safety of technical and technological facilities. In contrast to the well-known approaches to quality management and environmental safety of the states of technical and technological objects, which are not focused, as a rule, on the problem of accelerating calculations, the proposed generalized scheme of intelligent quality and safety management of the states of technical and technological objects is based on modified models of computing and technologies.

**Keywords:** intelligent control, technical and technological objects, computing models, environmental safety, technical condition.

## Введение

Для управления качеством состояний сложного технического объекта, необходимо провести оценку его технического состояния. Проведение оценки качества состояний сложных технических объектов требует использования показателей качества, т.е. количественные и качественные характеристики свойств объекта, входящих в его качество, рассматривая применительно к этапам разработки, производства или эксплуатации объекта сложной структуры.

Обеспечение эксплуатационной способности сложного технического объекта достигается за счёт управления его техническим состоянием. Однако, при управлении техническим состоянием сложного объекта возникает ряд трудностей. С одной стороны, многие современные сложные технические и технологические объекты отличаются отсутствием или недостаточной пригодностью к контролю технического состояния (диагностированию). С другой стороны, для определения технического состояния сложного объекта отсутствует общий системный подход при обработке и дальнейшем анализе разнородной информации – множества параметров, разного типа, описывающих его техническое состояние. При этом для промышленного технического оборудования, задача определения технического состояния часто является невозможной задачей. Это связано с невозможностью получения значений многих параметров, в процессе эксплуатации промышленного оборудования. Таким образом, процесс определения состояний сложных технических объектов для эффективного управления качеством, требует применения системного подхода.

Промышленные технические и технологические объекты, как правило, являются потенциальными источниками опасности, важным понятием для которых является «безопасность». При этом в последнее время на первое место выходят проблемы соответствия промышленного производства современным экологическим требованиям. Решение данных проблем невозможно без решения задачи обеспечения экологической безопасности состояния технических и технологических объектов. Данная задача прежде всего направлена на реализацию мероприятий, обеспечивающих допустимое воздействие факторов экологической опасности сложных технических объектов на окружающую среду и человека, что не может быть достигнуто без решения задачи утилизируемости объектов по истечении срока эксплуатации, выхода из строя или ликвидации.

В связи с этим представляет интерес создание системы классификации, которая позволила бы идентифицировать параметры, как качественные, так и количественные, которые определяют не только сам объект сложной структуры, его состав, но и состояние классифицируемого объекта, а также степень взаимодействия с окружающей средой (экобезопасность).

## Анализ проблемы

Основы функциональной систематики (ФС) базируются на концепции функциональных воздействий. Согласно данной концепции, процессы, реализуемые в материальном производстве, в экологии, в научных исследованиях, представляются как взаимодействие минимум двух материальных объектов, совершаемое в рамках определенной среды. Концепция функциональных воздействий реализуется в принципе применимости.[1]

Функциональные воздействия являются одной из основных форм движения материи и приняты в систематике в качестве основания деления технических процессов и объектов на иерархические системы – таксоны. Объекты, подвергающиеся воздействию или взаимодействующие с объектами-функционалами, условно названы исходными объектами. Исходными объектами определяется морфология служебных функций.

Функциональные взаимодействия технических объектов совершаются в рамках функциональной системы. Объектами системы являются (рис. 1) объект-функционал  $R_i$ , объект  $B_i$ , подвергающийся воздействию, и факторы внешней среды  $B_a$ .

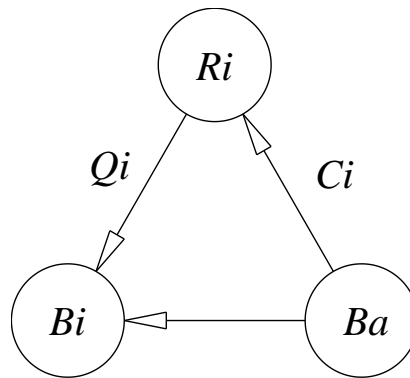


Рис. 1. Структура функциональных систем [2].

Объектами, определяющими функциональные воздействия  $R_i$  являются технические объекты, которые характеризуются целенаправленным воздействием на исходные объекты. Объекты, определяющие функциональные воздействия  $R_i$  подразделяются на вещественные объекты-функционалы  $M_i$ , а также поля, волны и микрочастицы-функционалы  $T_i$ , а также энергию-функционал  $J_i$  и процессы-функционалы  $P_i$ . [1]

Воздействия объекта, определяющего функциональные воздействия  $R_i$  на объект, подвергающийся воздействию, совершаются для реализации заданной служебной функции  $F_i$  (функциональная связь  $R_i-V_i$ ), что характеризуется логической записью:

$$R_i \rightarrow F_i \rightarrow V_i,$$

где  $R_i$  – объект, определяющий функциональные воздействия;  $F_i$  - служебная функция;  $V_i$  - объект, подвергающийся воздействию.

Функциональные связи  $V_a - R_i$  и  $V_a - V_i$  (рис.1) характеризуют совместимость объектов, определяющих функциональные воздействия с факторами внешней среды, которая характеризуется критерием совместимости  $C_i$ . Таким образом, определяется структура еще одной таксономической категории - нормы совместимости с факторами внешней среды:

$$C_i(V_a)$$

где  $C_i$  - критерии совместимости;  $V_a$  - материальные объекты-факторы внешней среды.

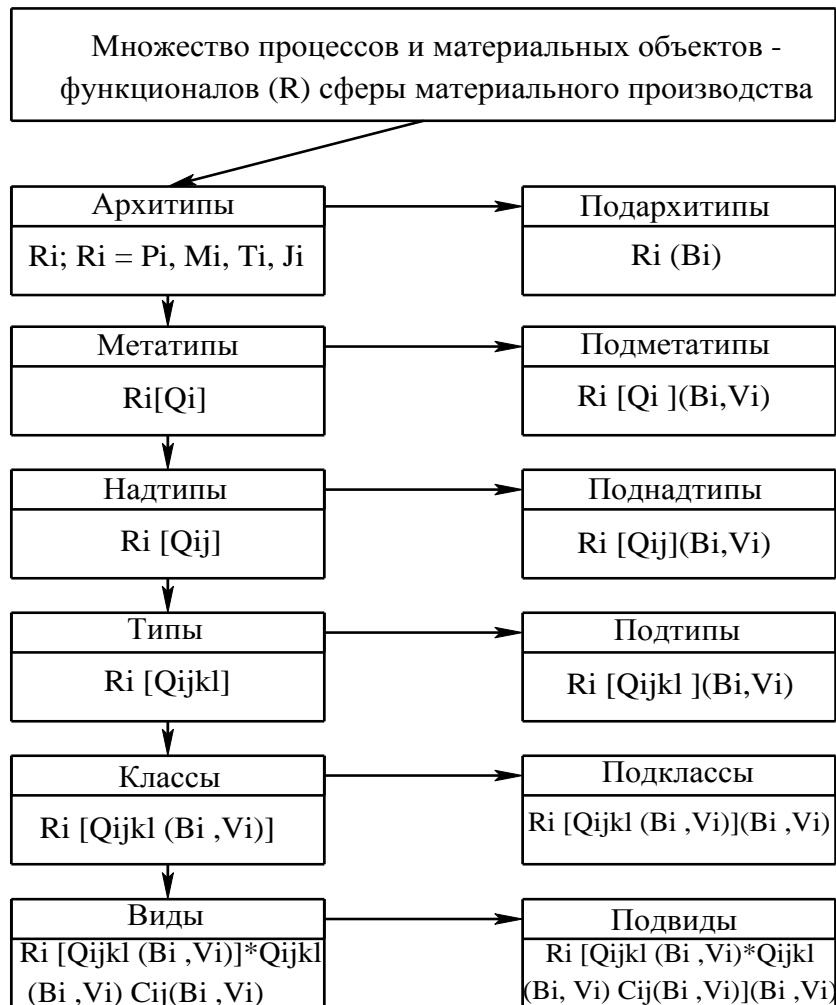
В зависимости от реализуемых функций технический объект может выступать или в качестве объекта-функционала, или в качестве объекта, на который направлено воздействие, или в качестве объекта внешней среды.

Воздействия или взаимодействия объектов-функционалов, исходных объектов и объектов-факторов внешней среды совершаются в процессах-функционалах  $P_i$ . Процессы-функционалы характеризуют выполнение служебных функций. Служебные функции  $F_i$ , исходные объекты  $V_i$ , нормы совместимости с факторами внешней среды  $C_i$  ( $V_a$ ) являются основаниями деления множества объектов-функционалов и процессов на соподчиненные иерархические группы – таксоны [2].

### Методика исследования

По общепринятым представлениям таксономия - это теория систематики. Принципы и основы декомпозиции в ФС характеризуются структурой функциональных формул (таксономической моделью). В таксономической модели (рис. 2) показана структура функциональных формул материаловедческой и машиноведческой систематики. В машиноведческую систематику включена и систематика технических объектов.

На рис.2 показаны:  $R_i$  - объекты и процессы, определяющие функциональные воздействия;  $Q_{ijkl}$  - служебные функции;  $(V_i, V_i)$  - мерон исходного объекта;  $(V_i, V_i)$  в позиции перед правой квадратной скобкой - мероны факторов внешней среды;  $(V_i, V_i)$  в позиции после правой квадратной скобки - мерон морфологического образа объекта, реализующего функциональные воздействия;  $C_{ij}$  - критерии совместимости с факторами внешней среды.



**Рис. 2. Таксономическая модель процессов и объектов сферы материального производства [3]**

Формулами определяется принадлежность систематизированного объекта и процесса к определенному таксону, а также способность объекта и процесса выполнять определенные служебные функции в заданных условиях взаимодействия с факторами внешней среды.

Таксономическая модель (рис. 2) является базовой моделью ФС. С некоторыми непринципиальными уточнениями, как показано ниже, модель используется в экологической систематике. Суть представленной модели заключается в последовательном удлинении формулы за счет приписки основания деления к делимому формулы. Например, формула класса материала M1  $[Q_{ijkl} (B_{ijkl})]$  состоит из формулы делимого типа M1  $[Q_{ijkl}]$  и формулы основания деления морфологии служебных функций  $B_{ijkl}$ .

Множество объектов и процессов-функционалов материальной сферы делится на следующие подмножества: P – процессы, определяющие функциональные воздействия; M – вещественные объекты, реализующие функциональные воздействия; T – поля, волны, излучения, микрочастицы и элементарные частицы-функционалы; J – энергия, определяющая функциональные воздействия.

Представленная таксономическая модель (рис. 2), предусматривает разделение множества функциональных процессов и объектов-функционалов на надтаксоны, таксоны и подтаксоны.

Для записи состояния объекта в формуле ФС модифицирована таксономическая модель, позволяющая записывать несколько параметров состояния объектах [4]:

$$R_i [Q_{ijkl} (B_i, V_i) * Q_{ijkl} (B_i, V_i) C_{ij} (B_i, V_i)] (B_i, V_i) (S_i (B_i, V_i), S_i (B_i, V_i), \dots) \quad (1)$$

где  $S$  – таксон состояния объекта  $i$  в  $S_i$  – нормированное значение параметра объекта в процентах ( $B_i, V_i$ ) после  $S_i$  – мерон параметра объекта

Например:

( $S98(L4.13.1)$ ,  $S100(L4.13.2)$ ,  $S86(L4.13.3)$ ) – нормированное значение габаритной высоты составляет 98%, нормированное значение габаритной ширины составляет 100%, нормированное значение габаритной длины составляет 86%. При этом предусмотрена возможность включения в модифицированную таксономическую модель комплексного показателя экобезопасности, который добавляется в формулу объекта-функционала, например:  $M6[Q1.12.4.1(B5.11.1)C1.5(A)](S_i(L6.1))$ , где ( $S_i(L6.1)$ ) – таксон состояния объекта, значение которого, характеризует комплексный показатель экобезопасности объекта  $L6.1$ . Показатель мерона  $S$ , указывается без запятой и всегда принимает трёхзначное значение.

Для получения оценки вектора состояний сложных технических объектов, состоящих из множества элементов, необходимо оценить техническое состояние каждого элемента системы в отдельности по параметрам, установленным для каждого элемента, затем рассчитать интегральные показатели вышестоящих по иерархии структур системы сложного технического объекта (рис. 3). В качестве такого показателя можно использовать индекс технического состояния (ИТС)[3], который также, как и комплексный показатель экобезопасности, может быть добавлен в виде таксона состояния объекта в формулу ФС.

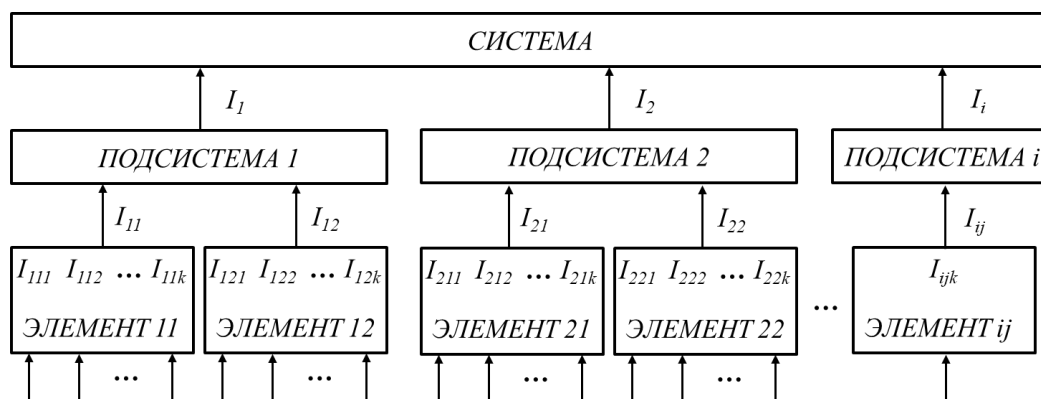


Рис. 3. Модель сложного технического объекта

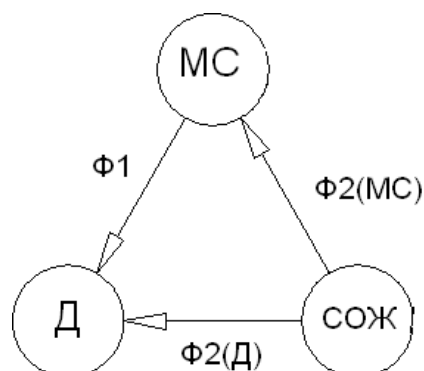
Использование модифицированных методов ФС для кодификации и поиска объектов-аналогов, как объектов-функционалов, объектов, подвергающихся воздействию и факторов внешних воздействий, позволяет организовать эффективный поиск эквивалентов. Безусловно, данный алгоритм осуществим, и работает при условии наличия базы систематизированных объектов. При этом один и тот же объект может быть указан в базе ФС неоднократно.

### Практическое приложение

Актуальной задачей является применение методов модифицированной ФС в машиностроении на примере металлообрабатывающего оборудования. Она включает в себя концептуальную основу, таксономию объектов и процессов - функционалов в сфере обработки материала, таксономии в сфере экологии, прежде всего смазывающе-охлаждающих жидкостей, структуры комплексов функциональных систем, образующих изучаемый объект, методов идентификации таких объектов, методологии параметрического моделирования, классификатора таксономических категорий и функциональной информатики.

Для металлорежущего оборудования в качестве объекта-функционала выступает металлорежущий станок (МС), объектом, подвергающимся воздействию – обрабатываемая деталь (Д), фактором внешней среды - смазывающе-охлаждающие жидкости (СОЖ) (рис. 4).

**Рис. 4. Структура функциональной системы металлорежущего станка**



Воздействия объекта-функционала МС на объект Д, подвергающийся воздействию (функциональная связь МС – Д), совершаются для реализации заданной служебной функции Ф1, что характеризуется следующей логической записью:

$$МС \rightarrow \Phi 1 \rightarrow Д \quad (2)$$

В данном примере служебной функцией является обработка детали. Функциональные связи Ф2(МС) и Ф2(Д) определяют совместимость объектов-функционалов с объектами-факторами внешней среды, в качестве которых может выступать смазывающе-охлаждающие жидкости (СОЖ). Совместимость характеризуется критерием совместимости К. Во множестве материальных объектов отсутствует жесткая прикрепленность к какой-то одной из рассмотренных позиций. В зависимости от реализуемых служебных функций материальный объект может выступать или в качестве объекта-функционала, или в качестве объекта, подвергающегося воздействию, или в качестве объекта-фактора внешней среды. Систематизируемый объект всегда рассматривается в качестве объекта-функционала - объекта, предназначенного для реализации служебной функции в определенных условиях взаимодействия с факторами внешней среды. Объекты, подвергающиеся воздействию или взаимодействующие с объектами-функционалами, являются исходными объектами, в данном случае это обрабатываемая деталь Д. Исходными объектами определяется морфология служебных функций.

Процесс воздействия или взаимодействия изучаемых объектов совершаются в процессах-функционалах П, реализующих служебные функции. Полученная структура является основой построения множества объектов-функционалов и процессов на иерархические группы – таксоны, которые объединены между собой общими признаками.

### **Выводы**

Существующие в настоящее время системы классификации (Универсальная десятичная классификация (УДК), Международная классификация изобретений (МКИ), библиотечно-библиографической классификации (ББК) и др.) неизбежно приводят к тенденции дифференциации знаний, что препятствует объективности в оценке применимости разных объектов и процессов материального производства и их взаимодействия. Основные недостатки существующих классификационных систем заключаются в том, что они не учитывают морфологические и функциональные особенности объекта, его способность выполнять определенные служебные функции в заданных

условиях взаимодействия с факторами внешней среды. Что в свою очередь, оказывает существенное влияние на эффективность поисковых механизмов в сфере материального производства.

Модифицированная ФС отличается от известных классификаций широтой классификационного пространства, ориентацией на использование информационных технологий, информационно-поисковым языком и рассчитана на длительную перспективу использования. Классификационное пространство систематики охватывает: материалы, продукты, машины, механизмы, оборудование, изделия, относящиеся ко всем областям знаний и отраслям промышленности; процессы, материалы и оборудование, необходимые для создания, испытания, изготовления и эксплуатации средств производства, транспорта, а также для развития науки; документы и носители информации.

Во многих областях машиностроения, например в атомной промышленности, остро стоит вопрос контроля технического состояния объектов. Создание системы мониторинга и хранения информации о техническом состоянии объекта не всегда возможно или требует больших трудозатрат. На основе модифицированных методов ФС возможно построение подобной системы, которая позволяет охватить всё пространство материальных и нематериальных объектов, что имеет особое значение при систематизации экологических факторов.

#### **Библиографический список**

1. **Манцеров С.А.** Развитие системы единой функциональной систематики для хранения данных о техническом состоянии объекта/ С.А. Манцеров, А.Ю.Панов // Вестник Нижегородского университета им. Н.И. Лобачевского. - 2013. - №6 - ч.1 - С. 235-238.
2. **Бреховских, С.М.** Функциональная компьютерная систематика материалов, машин, изделий и технологий, «Машиностроение» / С.М.Бреховских, А.П.Прасолов, В.Ф.Солинов // Москва, 1995. – 551 с.
3. **Манцеров С.А.** Создание баз данных объектов машиностроения на основе формул функциональной систематики // Вестник Воронежского гос. техн. ун-та. 2007. Т. 3. № 11. С. 171–176.
4. **Манцеров, С.А.** Управление техническим состоянием сложных систем на основе нечеткой модели / Е. А. Гаврилюк, С. А. Манцеров // Автоматизация процессов управления. – 2018. – № 1 (51). – С. 91–98.

## АВТОМАТИЗИРОВАННАЯ СИСТЕМА ПОДДЕРЖКИ ДОКУМЕНТООБОРОТА

Альшатер Махер Габер Али

ORCID: 0000-0001-7793-2891 e-mail: [althaere@mail.ru](mailto:althaere@mail.ru)

Нижегородский государственный технический университет им. Р.Е. Алексеева

В докладе на примере группы геодезических компаний ООО "ЗЕНИТ" рассматриваются основные процессы разработки информационной системы электронного документооборота. Описан алгоритм обработки входящей и исходящей документации. Приведена основная цель автоматизированной информационной системы - разработка подсистемы учета и контроля исполнения документов ООО "ЗЕНИТ". Описаны результаты, которые предполагается достичь при внедрении информационной системы на предприятии. Приведённые расчёты экономической эффективности внедрения информационной системы подтверждают целесообразность результатов работы. Срок окупаемости составил 138 рабочих дней, что составляет примерно 6,5 месяцев.

**Ключевые слова:** автоматизированная система, документооборот, документ, электронный.

## THE AUTOMATED SYSTEM FOR SUPPORTING THE CIRCULATION OF DOCUMENTS

Alshater Maher Gaber Ali

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

The article considers the main processes of developing an information system for electronic document management. The authors analyze the example of the group of geodetic companies LLC "ZENIT". An algorithm for processing incoming and outgoing documentation is described. The main goal of the automated information system is formulated as the development of a subsystem for accounting and control over the execution of documents by ZENIT LLC. The results that are expected to be achieved when implementing the information system in an enterprise are described. The results of calculations given in the article prove the economic efficiency of the implementation of the information system and confirm the feasibility of the results of the work. The payback period is 138 working days, which is approximately 6.5 months.

**Key words:** automated system, workflow, document, electronic.



## Введение

Для решения задачи оптимизации работы с управленческой документацией многие предприятия используют системы электронного документооборота, что позволяет сократить время, требуемое на подготовку и обработку документации, уменьшить затраты и исключить возможность появления ошибок. Все это дает для фирмы прямой и косвенный экономический эффект. Поэтому разработка информационной системы учета и контроля прохождения и исполнения документов на предприятии является актуальной задачей, особенно в эпоху цифровой экономики.

## Анализ

Проведенный анализ показал, что технология обработки информации на предприятии выглядит следующим образом:

Рассмотрим на примере группы геодезических компаний "ЗЕНИТ" процесс разработки информационной системы электронного документооборота.

Группа геодезических компаний ООО "ЗЕНИТ" является специализированной компанией, работающей на рынке геодезических технологий и геодезического оборудования. На предприятии принята смешанная форма организации работы с документами. В каждом структурном подразделении (отделах, службах) выполняются операции по созданию специализированных документов, их систематизации и предархивной обработке. Часть операций документооборота (прием и обработка поступающих и отправляемых документов, их регистрация, контроль за сроками исполнения, тиражирование, архивное хранение) сосредотачивается в центральной службе делопроизводства – административно- хозяйственном отделе (АХО).

В административно-хозяйственный отдел поступают входящие документы от сторонних организаций и исходящие документы от различных подразделений. Заведующий канцелярией регистрирует эти документы путем заполнения экранной формы регистрационной карточки. Автоматически при этом формируются журналы документов. После заполнения основных полей реквизитов входящие документы передаются руководителю, который наносит соответствующую резолюцию. После возвращения документов в отдел дополнительные данные об исполнителе и конкретном указании заносятся в базу данных. Остальная корреспонденция передается в подразделение или ответственным сотрудникам. Сведения об ответственном сотруднике вносятся в регистрационную карточку. Периодически отслеживаются сроки исполнения документов и своевременно принимаются меры, способствующие своевременному их исполнению. Также формируются отчеты и осуществляется поиск необходимых документов.

Перед разработкой и внедрением информационной системы электронного документооборота необходимо выполнить следующие мероприятия:

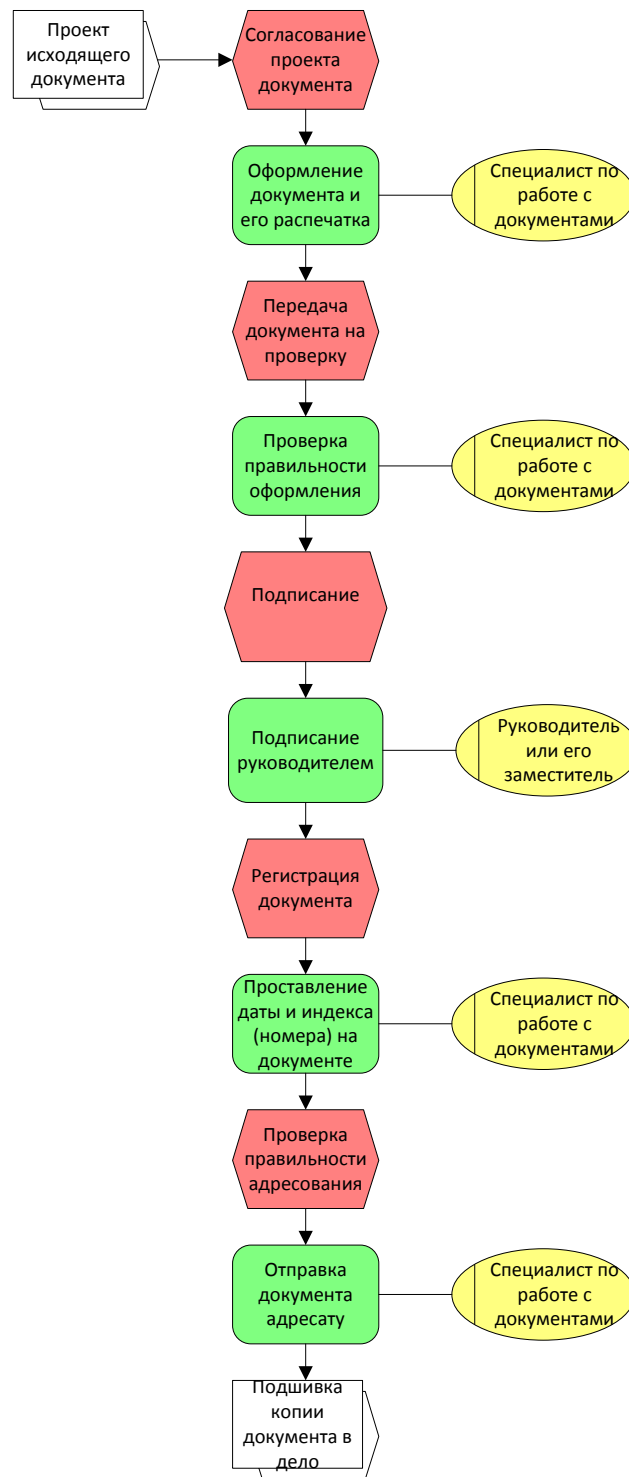
- 1) установить единый порядок работы с управленческими документами в соответствии с Государственной системой документационного обеспечения управления, в соответствии с ГОСТами и другими нормативами [1];
- 2) постоянно совершенствовать формы и методы работы с документами;
- 3) разработать и внедрить нормативные и методические документы по совершенствованию документационного обеспечения на предприятии;
- 4) осуществить контроль исполнения документов;
- 5) сократить документооборот, количество форм документов;
- 6) предусмотреть совместное участие в разработке и внедрении прогрессивных технологий документационного обеспечения управления на базе применения вычислительной и организационной техники всех заинтересованных лиц.

## Алгоритмы работы системы

Документы, поступающие на предприятие, проходят следующие этапы:



Рис. 1. Алгоритм обработки входящих документов



**Рис. 2. Алгоритм обработки исходящих документов**

Контроль исполнения является одной из главных функций управления предприятия. Цель контроля состоит в обеспечении своевременного и качественного исполнения поручений и решения вопросов. На предприятии организованы следующие типы контроля:

- 1) Контроль по существу решения вопроса.
- 2) Контроль за сроками исполнения задания.

Для регистрации в компьютере документа, поставленного на контроль, сотрудники АХО используют файл «Контроль и учет исполнения», который представляет собой документ (книгу) табличного процессора Microsoft Excel, в таблицу (лист) которого заносится следующая информация об исполнении документов:

Файл «контроль и учет исполнения документов»

№ п/п	Наименование документа	№ документа	Дата выдачи задания	Содержание задания	Срок исполнения работ	Дата фактического выполнения работ	Исполнитель	Отдел
1	2	3	4	5	6	7	8	9

Данный файл представляет собой обычную таблицу MS Excel, которая не имеет механизма эффективного поиска нужного документа, а также механизма отслеживания за сроками исполнения. Информация о поставленном на контроль документе просто заносится в таблицу. При снятии документа с контроля также заносится нужная информация. При этом, чтобы найти документ с просроченным сроком исполнения, приходится последовательно просматривать файл, что приводит к значительным потерям времени и снижению эффективности выполнения работы отделом.

Основная цель информационной системы - разработка подсистемы учета и контроля исполнения документов ООО "ЗЕНИТ".

Основными задачами, решаемыми подсистемой электронного документооборота, являются:

- 1) Автоматическая регистрация документа. Занесение в регистрационно-контрольную карточку сведений о документе.
- 2) Редактирование и удаление. Возможность внесения изменений или удаления документа.
- 3) Автоматическое формирование журналов регистрации документов на основе заполненной регистрационной карточки.
- 4) Автоматический поиск. Возможность быстрого поиска как обычного документа, так и находящегося на контроле по различным реквизитам.
- 5) Организация автоматического контроля за сроками исполнения документов [2].
- 6) Автоматическое формирование типовых отчетов.
- 7) Ведение и использование различных справочников.

При внедрении информационной системы на предприятие предполагается достичь следующих результатов:

1. Сокращение времени на обработку информации.
2. Увеличение объемов обработки и объемов хранения информации.
3. Уменьшение времени поиска документов и подготовки отчетов.
4. Уменьшение количества ошибок и повышение качества обработки документации.
5. Повышение технического уровня и коэффициента использования вычислительных средств.
6. Повышение исполнительской дисциплины.
7. Повышение информированности руководства за счет централизованной обработки информации, а также за счет повышения полноты и достоверности отчетов.
8. Уменьшение стоимости документационного обеспечения управления предприятием за счет перехода от бумажного делопроизводства к электронному, снижение стоимости копирования и передачи бумажных документов.

В системе должна быть предусмотрена возможность использования масштабируемых аппаратных платформ, в частности технологически простой замены серверов и устройств хранения информации на более производительные (емкие). Желательно, чтобы система документооборота могла поддерживать как пять, так, и пять тысяч пользователей, и ее способность наращивать мощность определялась только мощностью аппаратного обеспечения, на котором она установлена.

Выполнение этого требования может быть обеспечено с помощью поддержки промышленных серверов баз данных, таких компаний как Sybase, Microsoft, Oracle,

Informix, которые существуют практически на всех возможных программно-аппаратных платформах, обеспечивая тем самым максимально широкий спектр производительности.

Основные проблемы при работе с документами возникают в территориально-распределенных организациях, поэтому архитектура системы документооборота должна поддерживать взаимодействие распределенных площадок. Причем они могут быть объединены самыми разнообразными по скорости и качеству каналами связи. Для информационного обмена в системе должна использоваться корпоративная вычислительная сеть, в которую включаются как локальные, так и удаленные пользовательские системы. Система электронной документации должна взаимодействовать с электронной почтовой системой и автоматизированными системами управления предприятием.

Разработанная система состоит из отдельных модулей, интегрированных между собой, так как не всегда требуется внедрение всех компонентов документооборота, а только некоторый круг решаемых задач.

Следовательно, система документооборота должна строиться на основе открытых технологий, обеспечивающих ее дальнейшую модернизацию и развитие без переработки.

Также система должна иметь средства диагностики и индикации текущего состояния системы, используемых ресурсов; позволять наращивать количественные и расширять функциональные характеристики системы.

Система автоматизированного документооборота должна обеспечивать изменение своих характеристик, параметров, адресатов передачи документов и т.п. в зависимости от изменения регламента документооборота предприятия (норм, правил, порядка, форматов и т.п.).

Автоматизированные системы регистрации документов делятся на два класса. К первому классу относятся системы управления электронными архивами. Их основные функции: регистрация новых документов, хранение, поиск и извлечение документов с целью передачи в приложения, умеющие с ними работать. Ко второму классу относятся системы управления электронным документооборотом (СУЭД, DMS (Document Management System)). На них возложены функции управления документами на пути прохождения от одного пользователя - должностного лица к другому с возможностью контроля за их перемещением с фиксацией всех изменений и сопровождающих резолюций. В готовых системах одного класса могут встречаться функции другого класса.

На сегодняшний день в России существует множество систем электронного документооборота, призванных для автоматизации делопроизводства и административного управления предприятием.

Основными системами являются:

1. «Евфрат-Документооборот».
- 2 Система "Флагман": «Документооборот».
- 3 «Дело».
4. «1С:Предприятие»

Для разработки автоматизированной информационной системы документооборота был выбран язык программирования Delphi 6.

Delphi - императивный, структурированный, объектно-ориентированный, высокоуровневый язык программирования со строгой статической типизацией переменных.

Структурная схема разработанного пакета содержит программные модули различных классов:

- 1). Управляющие модули, предназначенные для загрузки меню или передачи управления другому модулю.
- 2). Модули, связанные с вводом, хранением, обработкой, выдачей информации.
- 3). Модули, выполняющие служебные функции (проверка пароля, ведение календаря, и т.д.).

## Описание модулей системы

Класс модуля	Идентификатор модуля	Функции
1	Main	Загрузка главного меню
1	DataModule	Контейнер невидимых компонентов
2	RegVhod	Регистрационная карточка входящего документа
2	RegIshod	Регистрационная карточка исходящего документа
2	RegVn	Регистрационная карточка внутреннего документа
2	OrgSpr	Справочник организаций
2	OtdelSpr	Справочник подразделений
2	KadrSpr	Справочник по кадрам
2	TipDokSpr	Справочник видов документов
2	SearchVhod	Поиск входящего документа
2	SearchIsh	Поиск исходящего документа
2	SearchVn	Поиск внутреннего документа
2	KontrIsp	Поиск документов на контроле
2	Rep1	Отчет о входящих документах за период
2	Rep2	Отчет о выполнении документов
2	Rep3	Отчет о документах, стоящих на контроле
3	Spravka	Дополнительная информация

## Результаты

Срок окупаемости составил 138 рабочих дней, что составляет примерно 6,5 месяцев.

Так как документ является основным способом представления информации на любом предприятии. Неоспорима важность сохранности и умелого использования информационных ресурсов в организации. Эффективность управления предприятием зависит от того, насколько разумно в нем организовано управление документооборотом.

Приведённые расчёты экономической эффективности внедрения информационной системы подтверждают целесообразность результатов работы.

В работе была достигнута поставленная цель и решены следующие задачи:

1. Произведен анализ существующего технического и программного обеспечения.
2. Получено обоснование проектных решений по видам обеспечения: информационное, техническое, программное и технологическое обеспечения задачи.
3. Обоснована экономическая эффективность проекта.
4. Разработана информационная система на языке программирования Delphi.

Разработанная информационная система была передана для внедрения ООО «Геодезическая компания «ЗЕНИТ».

## Библиографический список

1. ГОСТ Р 7.0.97-2016 «Система стандартов по информации, библиотечному и издательскому делу. Организационно-распорядительная документация. Требования к оформлению документов». — URL <http://protect.gost.ru/v.aspx?control=8&baseC=-1&page=0&month=-1&year=-1&search=&RegNum=1&DocOnPageCount=15&id=197990>. (Дата обращения 22.01.2022)
2. **Нагаева, И. А.** Программирование: Delphi : учебное пособие для вузов / И. А. Нагаева, И. А. Кузнецов ; под редакцией И. А. Нагаевой. — Москва : Издательство Юрайт, 2022. — 302 с. — (Высшее образование). — ISBN 978-5-534-07098-9. — Текст : электронный // Образовательная платформа Юрайт [сайт]. — URL: <https://urait.ru/bcode/493669> (дата обращения: 22.01.2022).

## АНАЛИЗ ВЕДУЩИХ ERP-СИСТЕМ НА РОССИЙСКОМ РЫНКЕ

Л.В. Кулагина

ORCID: 0000-0001-5734-411X e-mail: [kulaginal@mail.ru](mailto:kulaginal@mail.ru)

И.В. Козин

e-mail: [kozin@nttu.ru](mailto:kozin@nttu.ru)

А.М. Лабаев

e-mail: [labaev@nttu.ru](mailto:labaev@nttu.ru)

Нижегородский государственный технический университет им. Р.Е. Алексева

В докладе проведен анализ основных ERP-систем крупных предприятий, представленных на российском рынке с 2019 по настоящее время. Выделены показатели эффективности основных ERP-систем крупных предприятий. Описывается направление развития ERP-систем. Автор использует сравнительный метод для определения оптимальной ERP-системы для крупных компаний. Предложен единый подход к выбору ERP-системы крупных предприятий представленных. В заключение показано, что из проанализированного набора ERP-систем есть системы, подходящие для крупных компаний. В качестве заключения автор рассматривает возможный курс развития ERP-систем с учетом современных вызовов.

**Ключевые слова:** ERP-системы, анализ ERP-систем.

## ANALYSIS OF THE LEADING ERP SYSTEMS ON THE RUSSIAN MARKET

L.V. Kulagina, I.V. Kozin, A.M. Labaev

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

**Purpose:** The author analyzes the ERP systems of large enterprises represented on the Russian market.

**Design / methodology / approach:** The article selects the performance indicators of the main ERP systems of large enterprises. The article also describes the direction of development of ERP systems. The author uses comparative methods to determine the optimal ERP system for large companies.

**Findings:** Choosing the optimal ERP system for large enterprises.

**Originality/value:** The value of the whole article lies in the unified approach applied to ERP systems of large enterprises presented on the Russian market from 2019 to the present. In conclusion, the article shows that from the analyzed set of ERP systems there are systems suitable for large companies. As a conclusion, the author considers a possible course of development of ERP systems taking into account modern challenges.

**Key words:** ERP systems, analysis of ERP systems.

## Введение

Историческое появление ERP-систем можно отнести к 1960-м годам. В то время появляется система планирования потребностей в материалах или MRP-системы (MRP — MAterialRequirementsPlanning). В основе работы таких систем это лежит разузлования спецификации. На этом основании работа предприятия имеет актуальный вид. Агрегированную информацию из узлов использовали для планирования закупок и определения наполненности склада. Зародился рынок услуг по внедрению ERP-систем, в котором участвовали крупные компании, в 2000-е годы произошла консолидация поставщиков, появилось значительное количество ERP-систем для малого и среднего бизнеса, наиболее известными поставщиками, которых стали SageGroup и Microsoft [1]. В 1990-х годов появилось шесть наиболее успешных тиражируемых ERP-систем для крупных организаций.

К ним можно отнести следующие компании:

Baan – основными преимуществами является открытость среды ко многим базам данных. Недостатком является то, что данная система не является достаточно надежной, а также включает в себя сложный и долгий процесс внедрения.

Oracle – объединяет сервера, обслуживающие СУБД в большую базу данных, это позволяет увеличить производительность и уменьшить отказы. На каждом этапе свой шаблон документов. В связи с этим возникает сложность при работе с большим объемом документов.

PeopleSoft – построена на базе интернет-архитектуры, что позволяет объединить людей и архитектуру с помощью Web-браузера. Данная система имеет большую вероятность взломов. Безопасность данной системы вызывает вопросы.

SAP – основное преимущество в обеспечении прозрачности деятельности предприятия. Помогает крупным компаниям стандартизировать управление. Главным недостатком являются высокие требования к уровню подготовки персонала.

JD Edwards – предназначено для работы в условиях рынка, имеет гибкую структуру [2].

1С - флагманское решение «1С» для комплексной автоматизации средних и крупных предприятий самых разных сфер деятельности. Программа особенно актуальна для многопрофильных предприятий, с технически сложным производством и числом рабочих мест более десяти тысяч.

К 2021 году лидерами рынка являются компании SAP и «1С». В то же время доля Microsoft и Oracle, по его данным, существенно сокращается, однако SAP пока удается сохранить свои позиции на рынке в силу большей значимости российского рынка для вендора и наличия большой базы крупных клиентов. Согласно прогнозу AlliedMarketReseArch (AMR) мировой рынок ERP-систем к концу 2022 году должен был достигнуть 41,69 млрд долларов, при этом среднегодовой темп роста в течение 2019-2023 годов будет составлять порядка 7,2% [3].

С учетом импортозамещения, спрос на отечественные программные продукты класса ERP должен увеличиваться. Некоторые компании, как мелкие, так и крупные, должны будут изменить политику в вопросе развития ИТ-инфраструктуры с использованием иностранного программного обеспечения. «В связи с колебанием курса валют и резким удорожанием проектов, они начали тестирование российских продуктов, чтобы принять решение о полном или частичном переходе к их использованию», — говорит представитель корпорации «Галактика» Дмитрий Черных [4].

С учетом современных реалий происходит переориентирование направлений развития. Стихийно возросшее использование облаков поспособствовало быстрой адаптации к новым условиям: большим объемам видеоконференцсвязи, новым форматам ведения дел. С помощью широко развивающихся облаков стало возможным обеспечивать



такой объем проходящего трафика должной поддержкой и хранением без усилий от клиентов. В том числе это относится к ERP.

С 2020 года клиенты сами стали обращаются с запросами на их облачное размещение. Ведь облачные технологии - наиболее эффективное средство решения проблем, вызванных кризисом коронавируса. Заметим, что компаниям нужны готовые решения, они не хотят тратить время и инвестиции на долгие внедрения и разработки ТЗ.

В связи с импортозамещением можно выделить три основных проблемы: ограниченность функционала в российских ERP, недостаточная их производительность при полной автоматизации всего производства (например, на задачах разузловки) и стоимость такой миграции.

Один из важных трендов, о которых говорят эксперты, – упрощение интеграции ERP-системы в единое целое. Раньше это была заказная разработка, теперь же новые версии платформ позволяют клиентам реализовать интеграцию, полагаясь только на свои ресурсы.

Стоит отметить, что большие интеграционные возможности современных ERP-платформ позволяют реализовывать проекты не в форме внедрения единой системы, а как совокупность частей, ряда микросервисов, а также многочисленных интеграций с другими компонентами системы предприятия (хранилища данных, аналитические системы, системы оптимизации и прогнозирования и т.д.). Причем всё меньшее значение имеет, находятся ли те или иные системы и сервисы, с которыми интегрируется ERP-решение, в облаке или работают локально. Но данный подход может навредить построению целостной, единой, системы предприятия. Стоит опасаться закупать разные части у разных фирм и затем их сращивать в одно.

Так как ERP-системы давно вошли в нашу жизнь, то появляются новые запросы клиентов. Например, предоставления контрагентам информации прямо из внутренней ERP-системы, используя «личный кабинет клиента». То есть заказчик получает доступ к информации о расчетах. Это снижает затраты и время на рядовые операции, а раньше было невозможно из-за долгой предварительной подготовки и проверки. Можно расширить набор показателей, который менеджмент может видеть в режиме online. Если раньше это были, в основном, отгрузки и выручка, то сейчас все чаще анализируют затраты.

### **Особенности ERP-систем**

В современном мире автоматизация превратилась в привычный процесс в промышленности. Для принятия решения и анализа, информация передается на компьютер с помощью контроллеров, которые получили возможность управлять технологическими процессами на современных предприятиях всех отраслей промышленности. Поскольку автоматизация управленческих решений с помощью контроллера не предоставляется возможным, были разработаны интегрированные системы (среди них и SAP), которые являются конструктором взаимосвязанных модулей, созданных для управления производством.

В Германии в 1972 году, пятью сотрудниками компании IBM была основана компания SAP (Системный Анализ и Разработка программ). Главной целью была разработка стандартного программного обеспечения, которое объединяет в себе бизнес-процессы на предприятии в реальном времени. По прошествии двадцати пяти лет, компания SAP становится лидером в автоматизации по планированию ресурсами предприятия (Enterprise Resource Planning – ERP). На сегодняшний день она имеет филиалы и дочерние офисы во всех развитых странах.

В России компания существует 20 лет. В 1992 году в Москве открылся первый офис [24]. На сегодняшний день, компания имеет множество офисов в различных городах: Санкт-Петербург, Новосибирск, Екатеринбург, Ростове-на-Дону; Основным языком

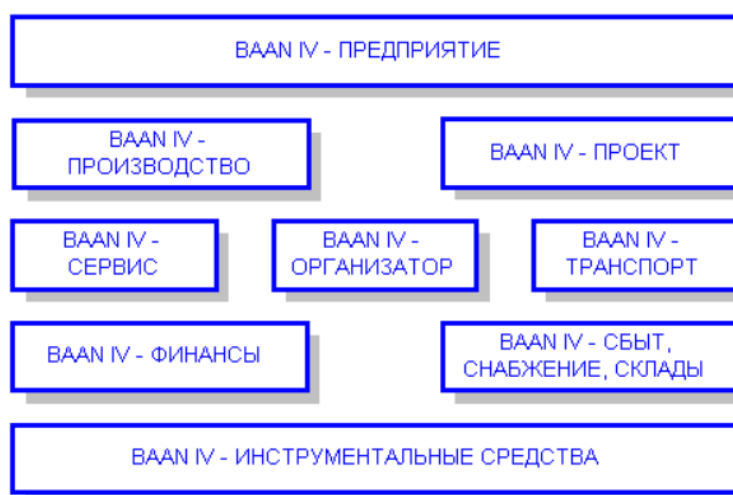
программных продуктов SAP является русский язык. Работой по внедрению SAP увлечены более 1000 русских, которые занимаются внедрением.

SAP ERP – информационная корпоративная система, в основании которой лежит методология ERP (планирования ресурсами предприятия), основным направлением является достижение оптимальности бизнес-процесса.

Основным плюсом SAP ERP системы является то, что данная система дает возможность руководству увидеть процесс производства в реальном времени, не вникая, при этом, в суть проблем, а также правильно и целостно оценить динамику движения процессов на предприятии, исключив при этом ошибки, которые могут быть связаны с человеческим фактором.

В современном условии конкурентной борьбы среди различных компаний, наиболее важным и значимым условием положительного результата организации, является наличие интегрированной информационной системы, которая дает возможность для оптимизации бизнес-процесса в организации, снижении себестоимости продукции и услуг [1].

BAAN IV – представляет собой интегрированную автоматизированную систему, поддерживающую такие направления бизнеса, как финансы, производство, сбыт, снабжение, склады, транспортные перевозки, сервисное обслуживание и проектно-конструкторские работы. Блоки, представляющие систему BAAN IV (рис.1).



**Рис.1. Основа системы BAANIV**

BAAN IV Tools - Программное инструментальное средство. Основное его предназначение – это совершенствование приложений и создание новых. Содержит набор программных инструментальных средств, которые предназначены для усовершенствования, документирования, перевода и поддержки программного обеспечения. Подсистема включает в себя модули: Бюджет проекта, Изделие, Классификация продукта, Конструкторские данные, Конфигурация продукция, Общие данные, Основной производственный план, Потребности в материалах, Потребности в мощностях, Серийное производство, Система управления качеством, Спецификация изделия, Технологический маршрут, Учет времени, Учет затрат, Цеховое управление.

BAAN IV Distribution – подсистема сбыт, снабжение, склады. Содержит модуль управления контрактам, товарно-материальным запасом и складским хозяйством, модуль управления партиями изделий и их отслеживания, модуль электронного обмена данными. Кроме того, данная подсистема дает возможность планировать потребности в материальных потоках. Подсистема включает в себя следующие модули: Изделие,

Маркетинг и продажи, Общие данные, Пополнение запасов, Потребности распределения, Управление закупками, Управление запасами, Управление комиссионными скидками, Управление партиями, Управление продажами, Управление хранением, Учет затрат, Электронный обмен данными.

BAAN IV Service - Подсистема «Сервис» включает в себя автоматизацию процесса сервисного обслуживания и ремонта, также, включает в модуль управления периодическим обслуживанием и текущим ремонтом, учет заявок, управление договорами, разработка графика работ для специалистов и графика обслуживания, анализ затрат. Подсистема содержит модуль: Основные данные, Управление установками, Контракты, Заказы на обслуживание, Анализ обслуживания.

FinAnce – финансы. Дает возможность работы с главной книгой, счетами кредиторов и дебиторов, осуществление контроля и регулирования операций, использование электронных методов учета поступлений и платежей. Подсистема включает в себя усовершенствованную систему финансового планирования с помощью использования обращений к первичной информации, модуль бухгалтерского учета и основных средств, отчетность и группировку компании. Подсистема состоит из модулей главной книги; общих данных; счет кредитора; распределение затрат; счета дебитора; управления денежными средствами; основных средств; систему финансового плана.

BAAN IV TrAnspOrtAtion – транспорт. Дает возможность автоматизирования внешних экспедиторских и транспортных услуг в рамках специализированной транспортной компании. Содержит следующие модули: управление заказам на хранение и транспортировку; учет товарно-материальных запасов на автотранспорт и горюче-смазочные материалы. Подсистема позволяет рассчитать стоимость фрахта, планирование погрузочно-разгрузочных работ и потребность в материальных потоках.

BAAN IV Project – проект. Позволяет управлять несколькими проектами и оценивать их. Подсистема включает в себя модуль планирования работ над проектами, разбитых по этапам, модуль который контролирует ход выполнения проекта в отношении плановых и оценочных показателей, при этом, позволяющий разработать прогноз, приготовить счет-фактуру и вычислить доход от проекта. Подсистема содержит следующие модули: Данные основные; оценка, бюджет проекта; потребность проекта; фактурирование и мониторинг проекта.

BAAN IV OrgAniser – организатор. Дает возможность быстрого, с малейшими усилиями внедрению любых продуктов данной серии (BAAN IV). Включает в себя модуль «Анализатора бизнес-потока» (BusinessFlowAnAlyser), мультипрограммного инструментального средства (MultimediAToolkit) для обучающих программ, информационной системы предприятия (EnterpriseInformAtionSystem).

Система ORACLE APPLICATIONS представляет собой высокопроизводительную многопользовательскую интегрированную систему учета, планирования и управления производством.

Система ORACLE APPLICATIONS состоит из основных комплексов:

Финансы - решает задачи управления основными фондами, финансового планирования, финансового анализа, консолидации, управления затратами, использования нескольких валют, управления денежными потоками. В комплексе финансы интегрируются и используются модули:

- Финансовый анализ (OracleFinancialanalyzer);
- Главная книга (OrAcleGenerAlLedger);
- Закупки (OrAclePurchAsing);
- Дебиторы (OrAcleReceivAbles);
- Кредиторы (OrAclePAyAbles);
- Основные фонды (OrAcleAssets).

Логистика – включает в себя задачу по управлению качеством (анализ качества продукта на стадии производства), планированию (система распределения планов, правила формирования, обработка, охватывающая все стадии жизненного цикла продукта: разработка, производство, реализация), управление снабжением (поддерживание бесперебойного снабжение предприятия сырьем и материалами), управлению материальными потоками (система инвентаризации и распределения материалов на предприятии), управление продажами, работы с клиентами. В комплексе используются следующие приложения:

- Планирование (OrAcleSupplyChAinPlAnning);
- MAsterScheduling/MRP;
- Управление материальными запасами (OrAcleInventory);
- Закупки (OrAclePurchAsing);
- Кредиторы (OrAclePAyAbles);
- Входящие счета (OrAcleOrderEntry);
- Товарная спецификация (OrAcleProductConfigurAtions);
- Дебиторы (OrAcleReceivAbles);
- Обслуживание (OrAcleService);
- Качество (OrAcleQuAlity).

Производство – представляет собой решение задач, связанных с разработкой новых продуктов, планированию и моделированию, управлению материальными потоками, управления себестоимостью, управления качеством.

В комплексе «Производство» используются приложения:

- Машиностроение (OrAcleEngineering);
- Структура продукта (OrAcleProductConfigurAtor);
- Счета материалов (OrAcleBillsofMAteriAl);
- Планирование материальных потоков (OrAcleSupplyChAinPlAnning);
- OrAcleMAsterScheduling/MRP;
- Производительность (OrAcleCApAcity);
- Управление материальными потоками (OrAcleInventory);
- Поставщики (OrAcleSupplierScheduling);
- Закупки (OrAclePurchAsing);
- Незавершенное производство (OrAcleWorkinProcess);
- Себестоимость (OrAcleCostMAnAgement);
- Качество (OrAcleQuAlity);
- OrAcle GEMMS.

Проекты – представляют собой решение задач, которые возникают при управлении проектами (определение цены, определение бюджета), аккумуляции стоимости (определение стоимости проектов), нормированию и анализу трудоемкости проекта, определению источников для выполнения проекта, определение дохода и прибыли. В комплексе «Проекты» используются приложения:

- Определение стоимости проекта (OrAcleProjectCosting);
- Определение финансового результата от проектов (OrAcleProjectBilling);
- Нормирование и трудоемкость (OrAclePersonAlTimeAndExpense);
- Данные складского учета (OrAcleApplicAtionsDAWAarehouse).

Человеческие ресурсы – решают задачу поиска и найма на работу персонала, учету обучения и тренингу персонала, учету заработной платы, учету списочного состава сотрудников, организационного планирования. В комплексе «Человеческие ресурсы» используются приложения:

- Списочный состав (OrAclePAyrole);

- Человеческие ресурсы (OrAcleHumAnResources);
- Развитие человеческих ресурсов (OrAcleTrAiningAdministrAtion);
- Выплаты, зависящие от продаж (OrAcleSAlesCompensAtion).

Управление рынком (маркетинг) – включает в себя решение задач, связанных с поддержкой управления продажами, управления маркетингом, анализом и завоеванием рынка, оперативной поддержки решений. В комплексе «Управление рынком» используются приложения:

- Финансовый анализ (OrAcleFinAnciAlAnAlyser);
- Данные складского учета (OrAcleApplicAtionsDAtAWArehouse);
- Продажи и маркетинг (OrAcleSAlesAndMArketiNg);
- Выплаты зависящие от продаж (OrAcleSAlesCompensAtion);
- Приложения для internet (OrAcleApplicAtionsfortheWeb).

Каждый из модулей подходит для отдельного использования. Так как использован общий словарь данных, интеграция происходит корректная.

Основные модули интегрированного пакета OrAcleApplicAtions:

- Модуль Главная книга (OrAcleGenerAlLedger) - система управления финансами, включающая в себя средства, которые необходимы для контроля финансов, сбора данных, подготовки финансовой отчетности.
- Модуль Дебиторы (OrAcleReceivAbles) - система управления расчетов дебиторов.
- Модуль Кредиторы (OrAclePAyAbles) – представляет собой учет и контроль расчетов с кредиторами, а также, учет форм скидок поставщиков.
- Модуль Закупки (OrAclePurchAsing) – управление закупками, которое необходимо для корректной и быстрой обработки заявок, заказов на приобретение, запросов предложения и поступлений.
- Модуль Основные фонды (OrAcleAssets) – представляет собой систему управления активами, которая необходима для корректной организации учета имущества и оптимального выбора стратегии учета и налогообложения.
- Модуль Управление материальными потоками (OrAcleInventory) – управление материальными запасами в складском хозяйстве.
- Модуль Входящие счета (OrAcleOrderEntry) – система сбытовых подразделений (ввод и управление заказами).
- Модуль Движение денежных средств (OrAcleCAshMAngement) – система управления движением денежных средств (мультивалютная).

PeopleSoft — американская компания, разработчик тиражируемого прикладного программного обеспечения для организаций HRMS, ERP, CRM, владевшая также активами JD EdwArds, поглощённая в 2005 году корпорацией OrAcle за 10,3 млрд.

В 1996 году была сформирована система PeopleSoft 6, она является целостной ERP-системой от PeopleSoft, интеграция которой произошла за счет модуля управления человеческими ресурсами, финансами и производством. За короткий срок, данная система стала лидировать среди остальных ERP. В 1998 году компанией PricewAterhouseCoopers был введен общепотребительный оборот акроним BOPSE — BAAn, OrAcle, PeopleSoft, SAP, JD EdwArds — обозначающий пятерку основных поставщиков ERP, отмечая тем самым ERP-систему от PeopleSoft как одну из лидирующих [20]. До 2000 года для работы с данной системой требовалась установка ключевым пользователям специальной программы. Но с 2000 года, был изменен интерфейс и перевыпущена система (PureInternetArchitecture).

В 2003 году PeopleSoft купил компанию JD EdwArds за сумму равную 1,7 млрд. долларов. Стоит отметить, что в момент покупки, система JD EdwArds была

единственным поставщиком ERP-систем для средних компаний, которые были внедрены и локализованы в мировом масштабе.

В 2004 году, после покупки JD Edwards, компания стала занимать второе место на рынке ERP с долей 12 процентов (при общем объеме рынка \$23,6 млрд), опережая при этом Oracle и проигрывая SAP. Вскоре после приобретения, компанией PeopleSoft была сформирована линейка из трёх ERP-систем:

- PeopleSoftEnterprise — унаследованная система компании, ориентирована на крупные предприятия,
- PeopleSoftEnterpriseOne — бывшая JD EdwardsEnterpriseOne, ориентирована на средние предприятия,
- PeopleSoftWorld — унаследованная от JD Edwards ERP- для мейнфреймов.

Начиная с версии 9 (выпуск 2006 года, компания стала входить в состав корпорации Oracle) система стала подразделяться на блоки:

- AssetLifecycleManagement — управление основными фондами предприятия;
- CustomerRelationshipManagement — модули управления взаимоотношениями с клиентами;
- EnterprisePerformanceManagement — управления эффективностью предприятия;
- ApplicationsPortal — портал клиентского самообслуживания, для развёртывания в организациях, ориентированных на массовое обслуживание абонентов, клиентов, заказчиков;
- EnterpriseServiceAutomation — модули автоматизации управления сервисной и проектной деятельностью;
- FinancialManagement — блок управления финансами организации, включая модули для бухгалтерии, контроллинга, казначейского учёта;
- HumanCapitalManagement — управления человеческими ресурсами организации, потомки исторически первого продукта компании — PeopleSoft HRMS;
- SupplierRelationshipManagement — управление снабжением и взаимоотношениями с поставщиками;
- SupplyChainManagement — управление цепочками поставок.

Проанализировав наиболее популярные современные ERP-системы, а именно:

- Система SAP;
- Система BAAN IV;
- Система ORACLE APPLICATIONS;
- Система PeopleSoft,

1С позиционирует себя как инструмент для объединения бизнес-процессов организации в одной мощной системе. ERP-системы предназначены для хранения и обработки большинства критически важных для компании сведений, они позволяют объединить все информационные ресурсы предприятия в единый механизм. Можно выделить 1С ERP Управление холдингом (рис.2) и 1С ERP Управление предприятием. Последнее наиболее востребовано. В нем выделяют следующие модули:

- Управление производством. «1С:ERP» – это единственный продукт «1С», в котором есть полноценный производственный модуль. В системе вы сможете создать заказ на производство, отследить ход работы над заказами, вести учет полного цикла производства, формировать графики, работать с маршрутными листами и т.д.
- Мониторинг и анализ показателей деятельности предприятия. В конфигурацию встроена система целевых показателей — панель управления для менеджеров всех уровней. Вы сможете выявлять проблемные места в

работе, контролировать задачи, анализировать ключевые процессы, составлять отчеты

- Управление затратами и расчет себестоимости. Программа позволяет эффективно оценивать расходы ресурсов по направлениям деятельности, вести учет затрат, рассчитывать себестоимость продукции
- Управление финансами и бюджетирование. В программе вы сможете оценивать эффективность бизнес-моделей, анализировать финансовые перспективы и результаты деятельности. С помощью бюджетирования решается одна из основных задач управления предприятием – комплексная оценка эффективности используемых бизнес-моделей.
- Организация ремонтов. «1С:ERP» позволяет вести учет объектов эксплуатации, регистрировать дефекты на производстве, планировать ремонтные работы, формировать заказы на ремонт
- Управление продажами. В системе вы можете выставить коммерческое предложение, зафиксировать отгрузки товаров, организовать доставку, оформить возврат и др.
- Регламентированный учет. «1С:ERP» поддерживает: общую систему налогообложения (ОСНО), упрощенную систему налогообложения (УСН), единый налог на вмененный доход (ЕНВД).
- Управление взаимоотношениями с клиентами (CRM). «1С:ERP» позволяет: управлять сделками, регламентировать процессы продаж, планировать события, хранить полную информацию по контрагентам, анализировать сделки, обрабатывать претензии клиентов, оценивать эффективность менеджеров.
- Управление персоналом и расчет заработной платы. В системе удобно вести штатное расписание, графики работы и отпусков, учет рабочего времени, рассчитывать заработную плату, проводить взаиморасчеты с сотрудниками, формировать отчетность и др.
- Управление закупками. Программный продукт позволяет: создавать новые заказы поставщику, подбирать поставщиков и условия закупок, сформировать заказ, контролировать исполнение заказов, вести мониторинг цен, составлять график поставок, оформить возврат и др.
- Управление складом и запасами. «1С:ERP» позволяет вести учет товаров на уровне складских ячеек (адресное хранение), использовать различные стратегии отбора для оптимизации склада. В программе автоматизирован процесс размещения товаров в складские ячейки, отбора товаров из ячеек и перемещения товаров из одной ячейки в другую.
- Интеграция с «1С:Документооборот». Совместное использование решений экономит время и избавит от необходимости перехода из одной информационной базы в другую: в карточках будут содержаться гиперссылки для доступа к любым учетным данным: файлы, процессы, задачи и др.

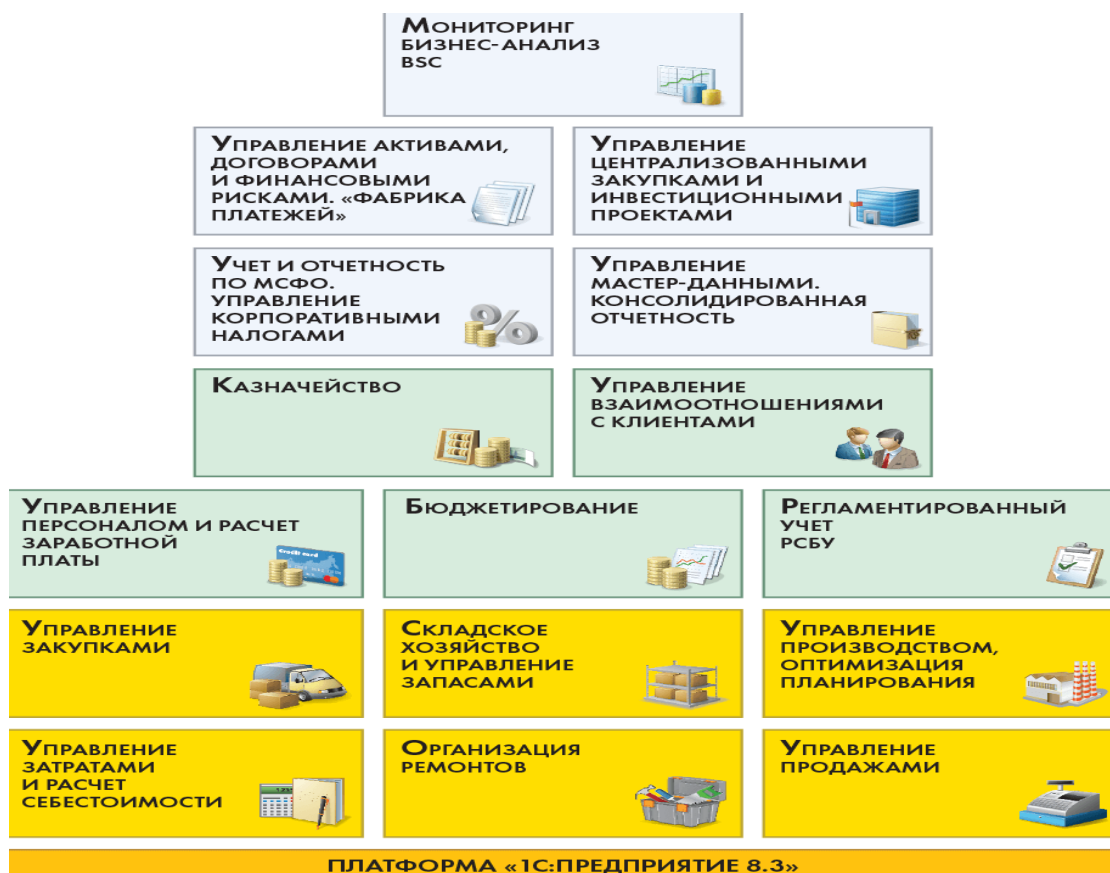


Рис. 2. 1С ERP Управление холдингом

### Выводы

Рассмотрев наиболее популярные ERP-системы, можно сделать следующие выводы:

- Наличие общей проблемы для всех систем при их внедрении, вне зависимости от выбора конкретной системы. Данной проблемой является подготовка данных, выгрузка и загрузка их в ERP- систему.
- При выборе системы, если учитывать импортозамещение, то в ближайшее время 1С станет самой востребованной ERP-системой.

### Библиографический список

1. Демьяненко, Ю.В., Добровинский, А.П.. Проблемы экономической оценки ИТ-инвестиций // «Инновации и инвестиции» - Томск, 2015 - с. 22-29
2. O'Leary, Daniel, L. Enterprise resourceplanning systems. - Cambridge University Press, 2014. - 232 с. - ISBN 0-521-79152-9.
3. RichArdsOn, Bruce. Five IdeAsAbout ERP. - ZDNet, 2015
4. Рудычева, Н., Зайнутдинов, Р. Первое настоящее сравнение «1С:ERP» и SAPERP. // CNews, интернет-портал. – М. 2018.
5. Дэниел, О'Лири. ERP-системы. Современное планирование и управление ресурсами предприятия. Выбор, внедрение, эксплуатация. - СПб: Вершина, 2018, 272 с. 8.
6. Питеркин, С.В., Оладов, Н.А., Исаев, Д.В. Точно вовремя для России. Практика применения ERP-систем.- М.: АльпинаБизнесБук, 2017, 368 с. 20.
7. Сигрид, Хагеман, Лиане, Вилл. SAPNetWeAver интервью, вопросы, ответы и пояснения: SAPNetWeAver сертификации Обзор. - М.: Лори, 2017,144.
8. Хетагуров, Я.А. Проектирование автоматизированных систем обработки информации и управления (АСОИУ). - М.: ОЛМА-пресс, 2016, 223 с. 23.



## ОСОБЕННОСТИ РАЗРАБОТКИ ИНСТРУМЕНТА ВЫГРУЗКИ ДАННЫХ ИЗ МОДЕРНИЗИРУЕМОЙ ERP-СИСТЕМЫ

Л.В. Кулагина

ORCID: 0000-0001-5734-411X e-mail: [kulaginal@mail.ru](mailto:kulaginal@mail.ru)

И.В. Козин

e-mail: [kozin@mntu.ru](mailto:kozin@mntu.ru)

Нижегородский государственный технический университет им. П.Е. Алексеева

В докладе рассмотрен процесс перехода от устаревшей ERP-системы к новой на примере перехода к системе SAP. Описываются этапы миграции данных. Рассмотрены особенности разработка инструмента выгрузки данных. Описаны преобразования структур хранения данных к нормальным формам. Приведены примеры данного процесса. Рассматриваются шаблоны загрузки данных.

**Ключевые слова:** ERP-системы, базы данных, корпоративные информационные системы.

## FEATURES OF THE DEVELOPMENT OF A TOOL FOR UNLOADING DATA FROM A HISTORICAL SYSTEM

L.V. Kulagina, I.V. Kozin

Nizhny Novgorod State Technical University, Nizhny Novgorod

**Purpose:** The paper considers the process of transition from a historical ERP system to a new one using the example of the transition to an SAP system. The stages of data migration are described. The features of the development of a data upload tool are considered. The paper describes transformations of data storage structures to normal forms. Data loading patterns are considered.

**Design / methodology / approach:** The article describes the features of the transition between ERP systems. The author reveals the features of uploading data from the historical system to the selected new one.

**Findings:** Features of database normalization.

**Originality/value:** The value of the entire article lies in the identified problems that employees will face during the transition from the historical system to the new one. In conclusion, the article shows that the problem of unloading data from a historical system to any other ERP system implies the same problem – data transfer, which will differ in correctness and accuracy

**Key words:** ERP systems, databases, corporate information systems.

## Введение

Система класса ERP (Enterprise Resource Planning — Управление ресурсами предприятия) — представляет собой корпоративную информационную систему (КИС), которая выполняет задачи, связанные с автоматизацией планирования, учета, контроля и анализом всех основных бизнес-процессов, и решением бизнес-задач в масштабе предприятия [1]. Основным плюсом является то, что департамент, работает с одной общей базой данных и им проще обмениваться между собой информацией. Вся информация хранится в единой базе данных, откуда она может быть в любое время получена по запросу.

Миграция данных представляет собой определенную завершающую последовательность работ, которые направлены на массовый одноразовый перенос из ранее использованной системы (система-источник) в новую (система-приемник). Следует отметить, что в исторической системе работа в данном случае прекращается. Интеграцию данных можно описать как постоянную часть архитектуры ИТ и отвечает за поток данных между разными системами и хранилищами данных. Также, интеграция данных – это процессом, с помощью которого происходит осуществление самого проекта. Схема миграции в общем случае выглядит следующим образом (рис. 1):

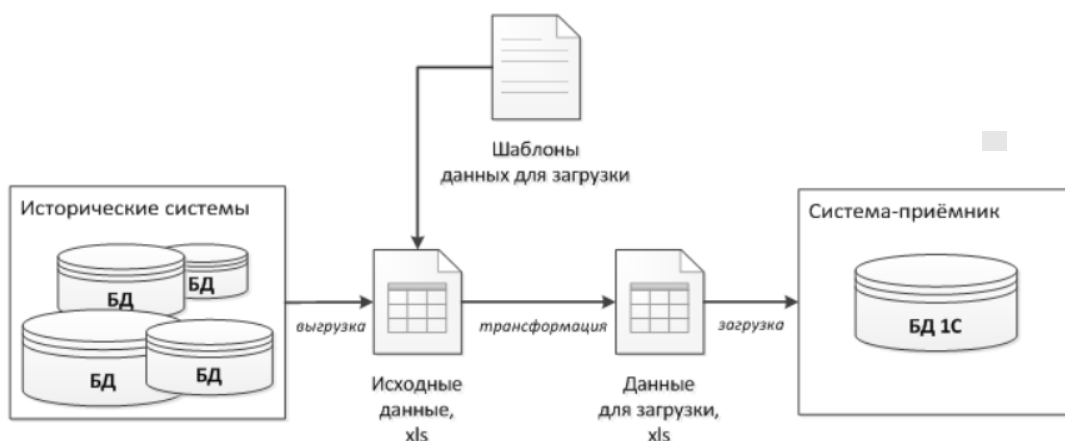


Рис.1. Схема миграции

Исторические системы – базы данных компании Заказчика, которые планируется полностью или частично заменить при внедрении новой системы.

Исходные данные представляют собой данные, выгруженные из исторических систем в произвольный формат xls-файлов. В данном случае формат xls представляется, как один из самых удобных, поскольку возможность выгрузки в xls-файл присутствует во многих учетных системах «предыдущих поколений».

Как альтернативу, возможно, рассматривать формат xml-файлов. Также существуют варианты использования промежуточной базы данных. Трансформация, конвертация – процесс преобразования исходных данных в данные для загрузки. Трансформация данных происходит в соответствии с шаблонами для загрузки. Результатом трансформации являются данные для загрузки. Данные для загрузки – данные, предназначенные для загрузки в систему-приёмник. Шаблоны данных для загрузки – описание таблиц данных для загрузки в целевую систему.

Можно выделить следующие этапы миграции.

К организационным этапам миграции можно отнести следующие пункты:

- Определение стратегии миграции. На данном этапе необходимо определить технологию проведения миграционных работ;
- Предварительный план миграции. План миграции по ходу проекта будет неоднократно корректироваться;

Периоды дат выгрузки данных из исторических систем, объемы данных. Периоды среза данных для миграций, даты тестовых и итоговой миграций. Данную информацию можно отнести к плану миграции;

Можно выделить типы данных, подлежащих миграции. Справочные данные, классификаторы, транзакционные данные, остатки, обороты и пр. Немаловажные вопросы связаны с проверкой качества, корректности и целостности данных в процессе миграции и по итогам. Стоит отметить, что актуальны моменты, связанные с откатом к предыдущему состоянию в случае сбоев.

Наиболее важным вопросом на первоначальном этапе являются технологические этапы миграции (рис. 2).



**Рис.2. Этапы миграции**

### **Подготовка шаблонов загрузки данных**

Шаблон загрузки включает в себя техническое описание таблиц данных, которые используются при загрузке, алгоритм и правило загрузки, предназначенного для имеющегося шаблона. Шаблон используется для нескольких связанных между собой таблиц в целевой системе-приемнике, или одной имеющейся таблице. В шаблоне должно быть:

- Описаны все поля xls-файла данных для загрузки, которые содержат: имя, признак, по которому определяется обязательность заполнения поля, пример, примечания.
- Описаны правила загрузки таблицы целевой системы, которые строятся на основании данных загрузки (очередность, алгоритмы поиска)
- Описано само заполнение полей таблицы целевой системы, если предусматривается переноса данных не идентичных друг другу файлу данных для загрузки (ссылочное поле).

Начинается с этапа подготовки шаблона загрузки данных. Происходит определение из какой системы в какую возможна выгрузка. Необходимо определить какие из данных понадобятся для дальнейшей работы. Определение всего списка, из которого будут браться данные занимает много времени и продолжается на всех последующих этапах. На данном этапе возможно выявление данных, которые находятся на печатных источниках, в данном случае их необходимо оцифровывать или вносить данные в таблицы со слов пользователей (ключевых). Чем больше будет выявлено и предоставлено нужной информации, тем корректнее будет конечный результат.

Стоит обратить внимание на Маппинг. Он представляет собой процесс сопоставления данных двух систем (система – источник и система - отправитель). Этап маппинга – является особо трудоемким этапом и может занимать больше пятидесяти процентов всей работы по миграции. В процессе маппинга данных необходимо выделить под этапы маппинга таблиц и маппинга полей. Маппинг таблиц, или маппинг шаблонов – сопоставление таблиц исходных данных и шаблонов данных для загрузки. Соответствие может быть как 1:1, так и N:N. В результате данной работы составляется и поддерживается реестр маппинга таблиц. Данный под этап необходим для следующего под этапом маппинга полей и для отслеживания общего состояния дел по маппингу. Основная проблема при переносе данных из исторической системы в систему SAP состоит в том, что в системе SAP существуют ряд полей, которые имеют одинаковые названия, но различные при этом значения (в зависимости от блока, в котором используются данные поля). Для детального анализа и создания алгоритма переноса данных из исторической системы в систему ERP SAP, необходимо более подробно изучить саму структуру системы SAP и блоки из которых она состоит.

Система SAP — представляет собой программное обеспечение, которое дает возможность автоматизирования профессиональной деятельности представителей разных специализаций. Такие приложения созданы для определенной отрасли и значительно упрощают работу в ней и связь с другими структурными единицами. SAP состоит из нескольких функциональных блоков, включает в себя управление, интеграцию и решение всех задач [3].

Система SAP R/3 представляет собой систему, состоящую из различных наборов прикладных модулей, которые необходимы для поддержки различных бизнес-процессов компании и имеют интеграцию в режиме реального времени.

Если рассматривать автоматизируемые бизнес-процессы, в таком случае, систему SAP можно классифицировать на компоненты и на несколько функциональных разделов таких как: финансы и бухгалтерия, контроллинг, логистика, управление персоналом, базис.

Отраслевые решенияДополнительные инструменты, при помощи которых автоматизируется бизнес.

При этом, функционал данных разделов реализуется модулями системы, а отношение какого-либо модуля к определенному разделу носит условный характер, так как модули глубоко интегрированы и часто одновременно реализуют различные функции.

FI – Бухгалтерия, представляет собой составной модуль, состоящий из разных подразделов: FI-AA – «Учет основных средств», FM – «Финансы», RCM – ведение

карточки договора). Решает задачи бухгалтерского учета предприятия. Включает в себя функции по отражению факта и планирования, данные Главной книги, Бухгалтерию дебиторов и кредиторов.

FI-AA – Учет основных средств. Модуль является глубоко интегрированным в блок FI, но также, данный блок является смежным с модулем PS «Система проектов», благодаря чему, становится возможным построение решения по управлению капитальным строительством.

FI-FM (PSM-FM) – Финансовый менеджмент, Бюджет, Управление финансами и т.д. Раньше этот модуль назывался FI-FM (FM – Fund MAnAgement), т.е. управление фондами. Потом его назвали PSM-FM (Public Sector MAnAgement – Fund MAnAgement).

FSCM – Управление цепочкой создания стоимости. Данный блок является относительно новым. Пользуется популярностью, поскольку дает возможность управлением финансами и рисками, реализуется учет ценных бумаг, кредитов и займов, депозитов. Также включает компоненты: Кредитный менеджмент (смежен с модулем SD «Сбыт»), Biller Direct, Управление наличностью и ликвидностью, Управление сборами, Dispute MAnAgement, Внутренняя наличность, Управление финансами и рисками (содержит функционал для сопровождения финансовых сделок).

IM – Управление инвестициями. Направленность на построение инвестиционной программы предприятия, работает в непрерывной цепочке с другими модулями (PS – система проектов, CO – контроллинг, PM – управление тех. обслуживанием и ремонтом оборудования). В настоящее время постепенно замещается продуктом RPM (Resource And Portfolio Management).

RE – Управление недвижимым имуществом. Позволяет управлять имуществом, от реестра, до ведения договоров аренды и т.д.

FI-BL – Банки. Модуль настроен на работу с банками.

Следующий раздел – «Контроллинг». Сам модуль CO – контроллинг. Решает вопросы по планированию и управлению экономикой предприятия. Имеют интеграцию во все модули системы.

PS – система проектов. Служит управляющим над проектами, в системе, как правило, используется для управления капитальным строительством или для управления уникальным производством.

PP – планирование производства. Рассчитанный на управление производственным процессом. Модуль имеет много расширений, которые выходят за рамки системы ERP, например, для балансирования мощностей. Функционал модуля решает задачи серийного, а также непрерывного производства.

PM – управление ТОРО (техническое обслуживание и ремонт оборудования). Планирование ремонтов, мониторинг оборудования (путем ввода в систему показаний контрольных приборов), калькулирования затрат на ремонты и т.д.

Раздел «Логистика» можно описать следующим образом.

MM – управление материальными потоками. Планирование потребности в материалах и услугах (в том числе, в интеграции с другими модулями: CO, PS, PM, PP и т.д.); приход на склад, перемещение, складской учет, фактурирование и т.д. Как и модуль FI, является самым распространённым и проработанным в части реализации требований менеджмента.

SD – управление сбытом. Решает задачи автоматизации продаж: ранжирование клиентов, предпродажная работа, обработка заказов клиентов, сопровождение поставок, фактурирование (подготовка документов: счета, накладные и т.д.).

Раздел «Управление персоналом (HCM)»

HCM-PA – Кадровый учет. Автоматизация всех основных процессов учета кадров.

OM – организационный менеджмент. Автоматизация управления структурой предприятия, ведение штатного расписания и т.д.

TM – Управление временными данными.

PY – расчет заработной платы. Имеет встроенный инструмент, который дает возможность написать скрипты и реализовывать расчет. Компоненты раздела рекомендуется установить на отдельно выделенном сервере. Это связано с тем, что в системе SAP ERP удобнее и проще реализация работы с полномочиями и штатным расписанием в блоке. Следовательно, актуальной является синхронизация систем ERP и HR, что в свою очередь, увеличивает стоимость блока.

В разделе «Базис» есть блок, который обеспечивает работоспособность, оптимизирует производительность, отвечает за полномочия, вносимые изменения в систему и развитие функциональности. Также, представляется возможным усовершенствования и разработки системы на языке ABAP/4 (на данном языке написан весь функционал). Он дает возможность реализации функций, не реализованных в стандарте системы, а также его модификацию.

Прежде чем внедрять ERP-систему, необходимо определить какой эффект принесет внедрение. Внедрение ERP-системы можно рассматривать как инвестиционный IT-проект.

Ниже рассматриваются основные из существующих способы оценки IT-проектов и их применимость к оценке эффективности внедрения ERP-систем.

Для определения экономической эффективности IT-проектов, как правило, используются те же финансовые методы, что и в оценке инвестиционных проектов. К данной группе методов относят: оценку чистой текущей стоимости (NPV), внутренней нормы доходности (IRR), срока окупаемости (PP) и т.п.

Но финансовых методов недостаточно. Вместе с традиционными финансовыми методами оценки эффективности IT-проекта, необходимо применять методы, позволяющие более полно учесть слабо формализуемые параметры и эффекты от внедрения IT-проектов. Для решения данной задачи были разработаны качественные и вероятностные методы оценки.

### **Базы данных для ERP систем, формы нормализации mssql**

Для переноса данных компании из исторической системы в систему ERP SAP, рекомендуется разработка инструмента выгрузки. Поскольку данные должны интегрировать в систему ERP SAP в полном объеме без потери какой-либо информации. Для разработки инструмента выгрузки необходимо использовать базы данных для ERP-систем – это реляционные БД и формы нормализации mssql.

Реляционная модель хранит любую информацию в таблицах в виде связанных записей с атрибутами. Необходимо наличие структур-таблиц. При этом атрибуты (столбцы) данной таблицы содержат разные типы данных [4]. Каждая запись в базе данных воспринимается как строка в таблице, а столбцы представляются атрибутами.

Отношения и типы данных представляют собой математические наборы, содержащие ряд атрибутов, которые в совокупности представляют собой базы данных и хранимую в ней информацию. Добавляя запись в таблицу, нужно распределить все её компоненты (атрибуты) по типам данных. Разные реляционные СУБД используют разные типы данных, и они не всегда взаимозаменяемы.

Подобные ограничения актуальны для реляционных СУБД, поскольку отношения между данными строятся на основе ограничений. В строке таблицы содержатся данные об одном объекте (товар), а столбцы таблицы содержат данные о характеристиках объектов — атрибутов (наименование). Сами строки таблицы должны быть представлены одинаковой структурой, состоящей из полей, которые содержат сами атрибуты объекта. В состав поля (столбца), входит какая-то одна характеристика объекта, при этом, она имеет определенный тип данных. В итоге, сами записи представляются одинаковыми полями, но отображение в них носит разный информационный характер. В реляционной базе данных каждая из таблиц имеет свой первичный ключ — он представляется полем или

комбинацией полей, с помощью которых возможно идентифицировать каждую из строк таблицы. В случае, когда в состав ключа входят несколько полей, он имеет название – составной. Ключ должен определять запись однозначно и иметь уникальность. Это нужно, чтобы предоставлялась возможность одной единственной записи по данному ключу. Для возможности упорядочивания хранимой информации в базе данных, также, возможно использование ключа.

Таблицы реляционной базы данных отвечают требованиям нормализации отношений. Нормализация отношений — представляет собой некие ограничения, связанные с форматированием таблиц, тем самым устраняется за двоение, непротиворечивости хранимых в базе данных, снижение трудозатрат на ведение самой базы данных. На сегодняшний день нет идеального четкого решения в области проектирования базы, поскольку один конечный вариант может подходить и одновременно не подходить к разным целям и бизнес-процессам. Во избежании потери времени и средств, рекомендуется не пренебрегать правилами работы с данными.

Нормализация баз данных представляет собой приведение структур хранения данных к нормальным формам (NF) [5]. Данных форм насчитывают восемь, но наиболее часто используемыми являются первые три:

Первая нормальная форма (1NF) - главное правило представляет собой неделимость значений каждого поля (столбца) строк – атомарность значений, следовательно, каждый из атрибутов содержит уникальное значение. Следовательно, не существует 1NF таблицы, в полях которых хранились бы списки значений. Помимо атомарности в первой нормальной форме должны соблюдаться правила:

Строки в таблицах не зависят друг от друга, следовательно, первая запись не влияет на вторую и обратно. Как будут размещены записи в самой таблице ни ан что не влияет.

Столбцы таблиц являются не зависимыми. На анализ информации влияния нет.

Каждая из строк является уникальной, для которой имеется первичный ключ, который состоит из множества или одного поля. Повторение первичного ключа невозможно, а, следовательно, является уникальным идентификатором.

Для приведения таблицы к первой нормальной форме, рекомендуется разбить таблицу на несколько отдельных (табл. 1, 2, 3).

**Таблица 1**

**Исходная ненормализованная таблица**

<b>Сотрудник</b>	<b>Номер телефона</b>
Иванов И.И.	000-11-22
	111-22-33
Петров П.П.	222-33-44

**Таблица 2**

**Таблица, приведенная к 1NF**

<b>Сотрудник</b>	<b>Номер телефона</b>
Иванов И.И.	000-11-22
Иванов И.И.	111-22-33
Петров П.П.	222-33-44

Вторая нормальная форма – отличительная особенность – это отсутствие зависимости не ключевых полей от части составного ключа. Таблица находится в 2NF, если она находится в первой нормальной форме, и при этом любой из её атрибутов, не входящий в состав первичного ключа, полностью зависит от первичного ключа. Это значит, что атрибут функционально зависит от всего первичного составного ключа,

но при этом не находится в функциональной зависимости от какой-либо из входящих в него атрибутов(частей).

Таблица 3

Приведение таблицы к 2NF

Начальник	Должность	Зарплата	Наличие компьютера
Иванов И.И.	Кладовщик	10000	Нет
Петров П.П.	Программист	30000	Есть
Петров П.П.	Кладовщик	20000	Нет

В табл. 3 представлены данные, в которых поля «Начальник» и «Должность» образуют первичный ключ. Зарплату сотруднику каждый начальник устанавливает сам, но её границы зависят от должности. Наличие же компьютера у сотрудника зависит только от должности, следовательно, зависимость от первичного ключа не полная. После преобразования данных к 2NF, образуется две таблицы (табл. 4 (а, б)).

Таблица 4(а)

После преобразования данных к 2NF

Начальник	Должность	Зарплата
Иванов И.И.	Кладовщик	10000
Петров П.П.	Программист	30000
Петров П.П.	Кладовщик	20000

В представленной таблице первичный ключ, как и в исходной таблице, составной, но единственный не входящий в него атрибут Зарплата зависит в данном случае от всего ключа (полно).

Таблица 4(б)

После преобразования данных к 2NF

Должность	Наличие компьютера
Кладовщик	Нет
Программист	Есть

Третья нормальная форма – практически не отличается от второй нормальной формы. Если вторая нормальная форма ликвидирует зависимости не ключевых полей от части ключа, то третья нормальная форма исключает зависимость не ключевых полей от других не ключевых полей. Таблица находится в 3NF, если она находится в 2NF, и при этом любой её не ключевой атрибут зависит (функционально) только от первичного ключа. Т. е. «нет зависимостей не ключевых атрибутов от других не ключевых атрибутов + 2 NF». Пример показан в табл. 5.

Таблица 5

Исходные данные для приведения к 3NF

Фамилия	Отдел	Телефон
Иванов И.И.	1	45-88-66
Сидоров С.С.	1	45-88-66
Петров П.П.	2	25-82-75

В результате приведения к третьей нормальной форме образуются две таблицы (табл. 6 (а, б)).



Таблица 6 (а)

## Приведенная к форме 3NF

Фамилия	Отдел
Иванов И.И.	1
Сидоров С.С.	1
Петров П.П.	2

Таблица 6 (б)

## Приведенная к форме 3NF

Отдел	Телефон
1	45-88-66
2	25-82-75

На практике, как правило, третья нормальная форма является достаточной и процесс проектирования реляционной базы данных завершается приведением к третьей нормальной форме.

## Выводы

Проблема выгрузки данных из исторической системы в любую другую ERP систему подразумевает одну и ту же проблему – перенос данных, который будет отличаться корректностью и точностью.

Процесс выгрузки данных из исторических систем может занять достаточное количество времени, особенно, если систем много, они разные и за них ответственны разные подразделения Заказчика. Необходимо учитывать данный момент при тестовых и итоговой миграциях. Исходя из этого можно сделать вывод, что актуальным и востребованным будет процесс разработки инструмента выгрузки из исторической системы в ERP-систему SAP.

Наиболее удобным вариантом представляется выгрузка в xls файлы. Многие старые IT-системы, в том числе ERP поддерживают такой вариант. Также могут быть варианты выгрузки в csv формат, dbf, xml форматы и прочие.

Стоит отметить, что по вопросам безопасности заказчик не всегда может предоставить выгрузки данных в полном объеме на этом этапе выгрузки исходных данных. Только саму структуру данных и несколько тестовых позиций. Следовательно, во время итоговой и тестовой загрузке может быть обнаружено некачественное введение данных в исходной таблице, что повлияет на итог загрузки, поскольку приведет к ошибкам.

Чтобы свести к самому минимуму проблему, рекомендуется предварительно проговорить объем тестовой выгрузки из системы-источника.

## Библиографический список

1. **Рид, Джон, Майкл, Доан.** Настольная книга SAP-консультанта. Книга, которая расскажет, как добиться успеха в мире SAP. - СПб.: Эксперт РП, 2016, 288 с.
2. **Питеркин, С.В. , Оладов, Н.А. , Исаев, Д.В.** Точно вовремя для России. Практика применения ERP-систем.- М.: Альпина Бизнес Бук, 2017, 368 с.
3. **Сигрид, Хагеман, Лиане, Вилл.** SAP Net We Aver интервью, вопросы, ответы и пояснения: SAP Net We Aver сертификации Обзор. - М.: Лори, 2017, 144 с.
4. **Сигрид, Хагеман, Лиане, Вилл.** SAP R/3. Системное администрирование. - М.: Лори, 2017, 480 с.
5. **Атре, Ш.** Структурный подход к организации баз данных. - М.: Финансы и статистика, 2015, 320 с.

## ОСОБЕННОСТИ ПЕРЕХОДА ОТ МОДЕРНИЗИРУЕМОЙ ERP К НОВОЙ

Л.В. Кулагина

ORCID: 0000-0001-5734-411X e-mail: [kulaginal@mail.ru](mailto:kulaginal@mail.ru)

А.М. Лабаев

e-mail: [labaev@nttu.ru](mailto:labaev@nttu.ru)

Нижегородский государственный технический университет им. Р.Е. Алексеева

В докладе рассмотрены особенности перехода от устаревшей ERP системы к новой на примере миграции в SAP. Определены основные моменты. Используются алгоритмы переноса данных, которые базируются на правилах сравнения типов данных. Учтены процессы переноса нормативно-справочной информации и данных. Показаны особенности переноса бухгалтерских данных, учета договоров и пользовательских настроек. В заключении особое внимание уделено проверке перенесенных данных из устаревшей системы в новую.

**Ключевые слова:** ERP-системы, перенос данных, перенос нормативно-справочной информации, переход от устаревшей системы к новой, система SAP.

## FEATURES OF THE TRANSITION FROM THE HISTORICAL ERP TO THE NEW SYSTEM

L.V. Kulagina, A.M. Labaev

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

**Purpose:** The paper considers the process of transition from a historical ERP system to a new one using the example of the transition to an SAP system. The stages of data migration are described. The features of the development of a data upload tool are considered. The paper describes transformations of data storage structures to normal forms. Data loading patterns are considered.

**Design / methodology / approach:** The article describes the features of the transition between ERP systems. The author reveals the features of uploading data from the historical system to the selected new one.

**Findings:** Features of database normalization.

**Originality/value:** The value of the entire article lies in the identified problems that employees will face during the transition from the historical system to the new one. In conclusion, the article shows that the problem of unloading data from a historical system to any other ERP system implies the same problem – data transfer, which will differ in correctness and accuracy

**Key words:** ERP systems, databases, corporate information systems.

## **Введение**

Проанализировав популярные на данный момент существующие ERP-системы, можно сделать вывод, что наиболее лояльной и удобной системой на момент 2021 года является SAP. Именно с системой SAP ERP начинает пользоваться такая крупная компания в России, как ПАО «Газпром». Момент перехода предприятия из исторической системы в систему ERP SAP является особо важным и трудозатратным. Это связано с тем, что возникают проблемы и сложности в сохранении большого объема данных, которое необходимо выгрузить и преобразовать для успешной загрузки в формате новой системы. Также, могут возникнуть сложности в выгрузке и переносе данных, которые связаны с конфиденциальностью выгружаемой информации, поскольку полный переход в систему ERP SAP подразумевает собой информацию о расчетных и банковских счетах, данные сотрудников предприятия, контрагентов, договоров, лимитов, заложенных на определенные отделы, размеры социальных льгот и выплат, которые предоставляет предприятие своим сотрудникам. В данном случае рекомендуется использовать специально разработанные шаблоны, которые необходимо корректно заполнить, так как именно от корректности заполнения данных шаблонов зависит успех переноса данных. В случае, если шаблон будет заполнен не верно, в систему ERP SAP информация загрузится не полностью или не являющаяся действительностью, а это означает, что данные сверки после загрузки не предоставится возможным произвести в полном объеме, следовательно, переход на новую систему не произойдет, что повлечет за собой колоссальные финансовые и временные потери для предприятия в целом. Именно поэтому к вопросу выгрузки данных из исторической системы следует относиться с максимальной сосредоточенностью и готовностью.

### **Алгоритмы преобразований данных**

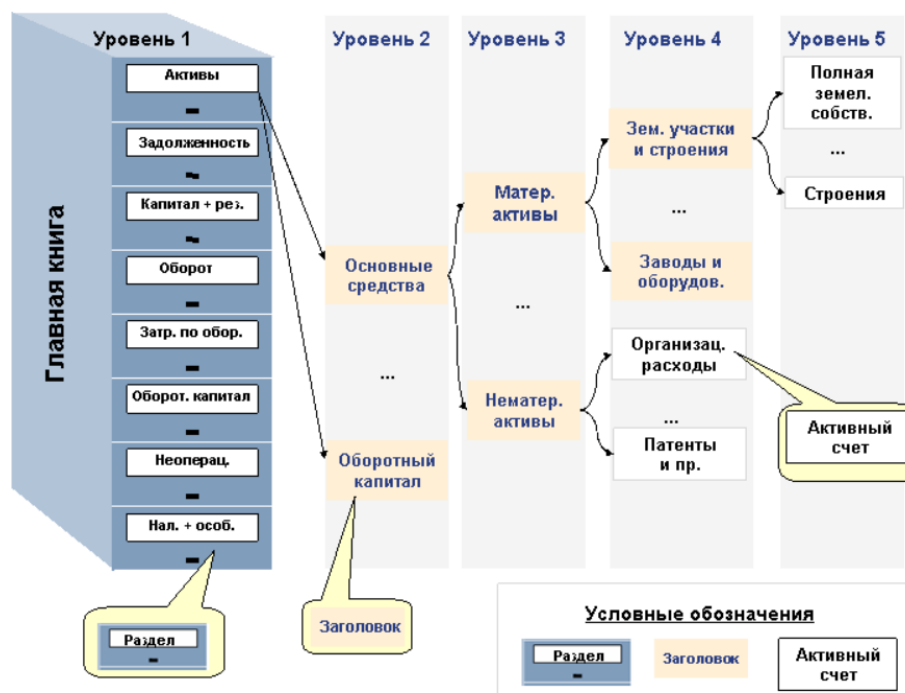
Первое с чего следует начать переход из исторической системы в систему ERP SAP – это выработка алгоритма преобразования данных. На большом предприятии в исторической системе используется определенный план счетов, который в любом случае будет отличаться от плана счетов в системе ERP SAP. Поскольку система SAP представляет собой свою специфическую настройку хранения данных. Основные объекты данных финансового модуля, их структура на уровне базы данных SAP

В данном случае для успешного перехода из исторической системы необходимо выгрузить и преобразовать план счетов, которые использует предприятие. В процессе внедрения консультанту необходимо использовать общепринятый принцип учета и государственное постановление, поскольку необходимо настройка бухгалтерского учета в соответствии с ними. После определения финансового (отчетного) периода создается план счетов, подходящий конкретной балансовой единице в системе ERP SAP. План счетов включает в себя все основных счета компании (счетов Главной книги), которые представляют собой основу отчетов и проводок по ним.

Счета в зависимости от раздела представляют подразделяются на синтетические и активные. Организация плана счетов следует общепринятому принципу учета, они подразделяются на активные, пассивные, счета собственного капитала, доходные, учитывающие себестоимость продаж, расходные, финансовые и др. (уровень 1 на рис.1). В каждом из разделов распределение счетов происходит согласно определенной логике по уровням, которая соответствует процессу финансового учета. Разделы – представляют собой организационные единицы верхнего уровня. Изменить предварительно определенные разделы, предоставляемые SAP, нельзя.

Каждый основной счет определяется как синтетический или как активный. Синтетический счет группирует все нижестоящие активные счета (уровень 2, 3, 4 на рис.1). В финансовых отчетах синтетический счет содержит суммарное сальдо всех нижестоящих активных счетов. Активные счета содержат проводки операций (уровень 4 и 5 на рис.1). Синтетические и активные счета различаются цветом текста:

синтетические обозначаются синим цветом, а активные - зеленым или черным. Зеленым цветом обозначены контрольные счета, определенные как счета главной книги, они устанавливаются по умолчанию для автоматической проводки операций в SAP Business One. Черным цветом выделяются активные счета, которые не выбираются как счета по умолчанию и устанавливаются консультантом (рис.1).



**Рис.1. Пример структуры плана счетов**

Главная книга в системе SAP ERP – это предоставление бухгалтерских книг, которая нужна для анализа компании со стороны финансов. Бухгалтерская книга включает в себя все финансовые проводки, создаваемые в Главной книге или других бухгалтерских книгах. Следовательно, главная книга является основным составляющим системы SAP ERP, так как именно она содержит в себе всю информацию о финансовых операциях компании. Главная книга включает в себя записи всех бухгалтерских операций, в ней можно посмотреть информацию о плане счетов, итоговые суммы на бухгалтерских счетах, данные оборотов, а также легко сформировать такие отчеты, как баланс предприятия и отчет о прибылях и убытках (рис.2).



**Рис.2. Новая главная книга системы SAP ERP**

Данные составляющие относятся к модулю FI в системе SAP ERP. Блок FI в системе включает в себя следующие возможности:

Позволяет управлять хозяйственными операциями с поставщиками на предприятии. Модуль охватывает спектр повседневных операций: управление записями кредитора и его задолженностью, отражение исходящих платежей, операций по закрытию периода. Кроме того, все документы, проведенные в бухгалтерии кредиторов (например, счета-фактуры и кредитовые авизо), автоматически отражаются на счетах Главной книги.

Учет расчетов с клиентами. Данное решение позволяет управлять основными записями клиента и дебиторской задолженностью, отражает входящие платежи и операции по закрытию периода. Все документы, проведенные в бухгалтерии дебиторов, автоматически отражаются на счетах Главной книги. Также существует система напоминаний для оповещения контрагентов о необходимости внесения оплаты.

Учет банковских операций позволяет создавать и обрабатывать исходящие и входящие платежи, кроме того, включает в себя обработку банка и работу, связанную с управлением наличностью.

Система позволяет работать одновременно в нескольких кассовых книгах, выполняя при этом проводки по счетам. Автоматизирован расчет начального и итогового сальдо, а также расход и приход, вводимый в кассу.

Оперативное закрытие финансового периода может быть настроена с помощью функциональности закрытия финансового года, месяца или дня. Есть возможность выполнения обмена валюты по курсам, возможно управлять периодом проводки, процедурой выравнивания счетов, переоценкой валюты. Также, есть возможность формирования отчетности.

Составление бухгалтерской отчетности в соответствии с требованиями законодательства РФ. В соответствии с требованиями ПБУ составляется бухгалтерская и налоговая отчетность, корреспонденция счетов, а также печатные формы (счёт-фактура, журналы счетов-фактур, накладная ТОРГ-12, пакет стандартных форм материального учёта).

Описанное выше показывает, насколько важным является определение плана счетов и составления маппинга счетов главной книги исторической системы и системы SAP ERP. Поскольку все хозяйственные операции в новой системе будут построены именно на счетах главной книги с определением корреспондирующего счета. И от правильности данного маппинга зависит успех проводимого перехода в новую систему SAP ERP.

Для успешного перехода в систему SAP ERP, необходимо после составления маппинга счетов, в исторической системе сформировать баланс на конец отчетного периода, которым будет выполнен перенос в систему SAP ERP, поскольку перенос данных лучше производить загрузкой остатков на конкретные счета, соответствующие счету SAP согласно маппингу.

После составления необходимых таблиц соответствия, рекомендуется выполнить преобразование данных с помощью SQL.

Преобразования данных являются контекстно-зависимыми. Необходимо указывать явные преобразования, обеспечиваемые функциями SQL. Это диктуется следующими соображениями:

- если применяются функции явного преобразования, SQL – операторы становятся проще для понимания;
- автоматическое преобразование данных может негативно влиять на производительность, особенно если столбец преобразовывается к типу данных константы;

- алгоритмы неявного преобразования могут меняться от версии к версии и от одного продукта ORACLE к другому и, следовательно, поведение явных преобразований более предсказуемо.

### Правила сравнения типов данных

Правила сравнения положительных чисел стандартные. В то время, как правило сравнения отрицательных чисел представляет собой гипотезу о том, что все отрицательные числа считаются меньше, чем положительные. Например, -2 меньше 200, а -300 меньше, чем -4.

Данные, которые состоят из символов придаются сравнению посредством сравнения каждого символа. Это происходит до выявления несовпадения определенного символа. При выявлении разницы в символах больше считается та строка, в которой имеется больший символ. В том случае, когда имеются все одинаковые символы в нескольких строках, то анализируется длина строк и меньшей считается та строка, в которая является короче.

Символы сравниваются в соответствии с «последовательностью слияния» компьютера. Один символ считается наибольшим в том случае, когда он представлен после него в слиянии последовательности. Следует отметить, что верхний и нижний регистры не считаются эквивалентными [6].

Таблица 1

#### Описание псевдостолбцов

Псевдостолбец	Возвращаемое значение
SequenceCURRVAL	Возвращает текущее значение указанной последовательности. Должно быть специфицировано <code>sequence_name</code> .
sequenceNEXTVAL	Возвращает следующее значение указанной последовательности. Должно быть специфицировано <code>sequence_name</code> .
ROWID	Возвращает ROWID для строки.
ROWNUM	Возвращает последовательный номер выдаваемой строки из таблицы или объединения строк. Первая выбранная строка имеет ROWNUM =1, вторая 2 и т.д.

Псевдостолбцы аналогичны столбцам таблицы, но в реальном времени в таблице не хранятся. Сами запросы могут брать информацию из псевдостолбцов. Псевдостолбцы нельзя добавить, изменить или удалить (табл.1).

Объекты базы данных

Объекты базы данных представляют собой элементы, которые поддерживаются SQL. Большое количество объектов занимает большой объем памяти в базе данных [4]. Существуют также объекты, которые несут в себе функцию хранения других объектов определенной базы данных, такими объектами являются табличные пространства. Уникальным объектом, хранящий в себе данные пользователя, к которому возможно обращение при запросе, являются таблицы.

Существуют правила, наименования объектов в базе данных [5]:

1. Имя объекта базы данных должно быть длиной от 1 до 30 символов (за исключением тех, объектов, чьи имена ограничиваются количеством символов равным 8).
2. Кавычки в имени не допускаются.
3. Имя не должно является регистрочувствительным
4. Название имени должно начинаться с буквенного обозначения.

5. Имя может содержать только символы A-Z, 0-9, \_, \$ и # (использование \$ и # специально оговаривается).

6. Имя не может дублировать используемые в программе слова ORACLE.

7. Имя не должно дублировать имени другого объекта базы данных, владельцем которого является один пользователь.

8. Некоторые из объектов базы данных могут в будущем, быть определены, вместе с именем пользователя, который является владельцем данного объекта.

9. Так же в имени может содержаться любая комбинация символов, при условии, что имя будет заключено в кавычки, тогда, правила 3-6, представленные выше, не действуют (в этом случае имя может содержать пробел).

Псевдонимы (алиасы) таблиц и столбцов, должны соответствовать данным правилам, несмотря на то, что не являются объектами базы.

При этом, важным условием является то, что должно быть согласование списка столбцов и значений. А также, не маловажным условием является то, что поле с первичным ключом (-ами) должно быть введено в строгом порядке.

Создание таблиц строиться в порядке подчиненности или в порядке связей между этими таблицами. Первыми начинают создаваться главные таблицы (родительские), только после этого создаются вспомогательные таблицы [2]. Оператор определения таблиц использует выражения, представленные ниже:

Вводить ограничения и связь на столбцы или на таблицу в целом возможно или после создания всех таблиц, или в момент создания. Для этого рекомендуется использовать команду ALTER TABLE.

С помощью выражения Alter table возможно изменить определение таблицы, а именно:

- 1) добавить столбцы или ограничения;
- 2) модифицировать определения существующих столбцов (например - тип данных или NOT NULL);
- 3) отменять ограничения;
- 4) модифицировать будущее распределение памяти;
- 5) отметить, будет выполняться backup для данной таблице.

### **Создание витрин данных для переноса НСИ и данных**

Витрина данных— хранилище данных (срез), который представляет собой массив тематической, узконаправленной информации, ориентированный на пользователей одной рабочей группы или департамента [1].

Первоначально для переноса данных нужно подготовить исходную информационную базу для переноса.

При витрин данных самым сложным и трудозатратным является получить четкий набор требований от клиента.

При проектировании витрин данных самым главным является - спроектировать эффективную структуру витрины. Поскольку, рядовая учетная система оптимизирована только для хранения информации, поля данных распределены по множеству таблиц. Именно поэтому поиск определенных полей, которые описывают конкретный запрос - это трудоемкий, медленный процесс, кроме того, он отрицательно воздействует на



**Рис.3. Схема для витрины данных продаж**

Оптимизация координированных витрин заключается в создании структуры, с помощью которой возможно оперативнее проанализировать данные и составить отчеты, необходимые для ведения предприятия. На мой взгляд, наилучшей в данном случае является схема "звезда".

В схеме звезда представлено два типа таблиц - таблицы фактов, которые содержат историю транзакций, в них также хранятся числовые показатели и идентификаторы (ID поля) для измерений и таблицы измерений [3].

В схеме "звезда" таблицы фактов окружены вспомогательными таблицами измерений, за счет этого, пользователь может быстро погрузиться в данные (drill down) для выявления корреляций между измерениями и элементами в таблицах фактов (рис.4).



**Рис.4. Таблицы фактов и измерений**

Формирование запросов представляет собой набора простых однонаправленных объединений (таблицы фактов и каждого из измерений). Главная сложность в создании витрин состоит в организации ключевых этапов:

извлечение данных из исходных систем (extract) – требует точного знания структуры исходной системы;

преобразование данных в нужную форму (transform) – необходимо, поскольку недостающие, неполные или неточные данные снижают качество анализа.

Процесс преобразования включает в себя множество функций, таких как:

- реструктурирование файлов данных, записей и полей;
- удаление избыточных данных;
- декодирование и трансляцию значений полей;



- повышение качества представления читаемых данных;
- проверку их достоверности;
- расчет новых значений для одного или нескольких исходных столбцов;
- упрощение данных и изменение их типов;
- последующая загрузка в целевую систему (load).

Для этого используется специальный ETL-инструмент (Extract, Transform, Load). После установки запуск идет по определенному расписанию.

Преобразованные данные загружаются в целевые витрины данных и становятся доступными пользователю для анализа и отчетности. Но главным условием является установка поддержки информации об измерениях для платформы доступа/хранения.

### **Перенос бухгалтерских данных**

Перед тем, как переносить бухгалтерские данные, необходимо проверить, чтобы все счета имели управление открытыми позициями.

Для счетов, которые ранее не имели управления открытыми позициями и которые будут так же использоваться в системе SAP, рекомендуется выполнять перенос сальдо нарастающим итогом в одной проводке. К таким счетам в основном относятся налоговые счета.

Проводка по налоговым счетам в системе SAP ERP осуществляется автоматически. Для этого следует поставить отметку в поле «только автоматическая проводка» в основной записи каждого счета. В данном случае ручная проводка запрещена. Но в случае, если необходимо выполнять ручные проводки по данным счетам, то данную отметку нужно установить только после переноса данных.

Для счетов, которые имели или имеют управление открытыми позициями в системе SAP, необходимо отдельно проводить открытые позиции. Это выполняется пакетным вводом или вручную.

Счета дебиторов и кредиторов; контрольные счета, перерасчетные счета и банковские субсчета

Счета дебиторов и кредиторов в системе SAP имеют управление открытыми позициями всегда. Для таких счетов необходимо проводить каждую открытую позицию пакетным вводом или вручную, поскольку проводка выравнивания отмечается отдельно создаваемым документом.

В системе SAP также используются некоторые специальные основные счета. К ним относятся контрольные счета, перерасчетные счета и банковские субсчета.

Контрольные счета используются для параллельной проводки тех бизнес-операций, которые система проводит по счетам вспомогательных книг (счета дебиторов и кредиторов), непосредственно в Главной книге. Чтобы обеспечить корректность проводок по контрольному счету, следует задать номер контрольного счета в основной записи дебитора или кредитора. Проводки по контрольным счетам производятся автоматически при переносе открытых позиций дебиторов и кредиторов. Следовательно, переносить сальдо контрольных счетов не нужно, если эти же счета использовались ранее в системе предприятия.

Перерасчетные счета и банковские субсчета должны иметь управление открытыми позициями. На момент переноса сальдо на таких счетах должно быть равно нулю, это означает, что проводки переноса не нужны.

Перед переносом старых данных из бухгалтерских документов, необходимо разделить счета на: счета с управлением открытыми позициями и счета без управления открытыми позициями.

Также рекомендуется создать счет переноса для корреспондирующей проводки. Данный счет должен подходить и для налогов, поскольку поле «налоговая категория» в основной записи должно быть пустым.

Налог автоматически рассчитывается во время проводки и проводится по налоговому счету. Это происходит при проводках по основным счетам, где поле «налоговая категория» в основной записи заполнено значением. Новые проводки налога переносить не следует, это необходимо чтобы не было за двоения, поскольку проводки налога переносятся вместе с сальдо.

Корректировка налога происходит в момент выравнивания платежей при платеже сконто. Настроено определение определенного кода налога из позиции ДТ или КТ или основного счета, но только в том случае, когда был использован один код налога.

В момент переноса старых данных происходит корреспондирующая проводка. Но данный счет для налога не допустим, в то время как отдельная позиция код налога не имеет. В момент переноса открытых позиций, рекомендуется вводить код налога в позиции дебета или кредита. В итоге счет переноса содержит нулевое сальдо после выполнения всех проводок. Далее рекомендуется блокировать счет, во избежании ошибочных проводок.

С помощью пакетного ввода необходимо перенести открытые позиции счетов кредиторов и дебиторов. При этом, система автоматически производит проводку по контрольным счетам дебиторской и кредиторской задолженности. При этом операции Особой главной книги и сальдо основных счетов необходимо проводить вручную (авансовые платежи).

Так же необходимо составить налоговую декларацию для того периода, который только что был закрыт в прежней системе.

Позиции дебиторов и кредиторов содержат условия платежа и базовую дату условий платежа (для расчета срока оплаты). Условия платежа из основной записи дебитора или кредитора можно перенести с помощью пакетного ввода или вручную.

Важным атрибутом бухгалтерского документа является код проводки. С помощью него определяется то, по какой стороне учёта будет проходить вводимая сумма.

### **Перенос учета договоров**

Карточка договора в системе SAP ERP представляет собой отдельный модуль, который называется RCM.

В системе SAP, карточка договора входит в блок RCM. Сама карточка договора состоит из нескольких разделов, каждый из которых несет в себе определенную информацию о договоре. Рассмотрим более детально основные из разделов.

Вкладка «общие данные» содержит данные о виде, категории договора, кураторе, номер карточки договора в прежней системе, предмет договора, дата создания карточки договора, дата начала действия, дата заключения, дата окончания действия, дата окончания платежей, информация о статусе. Данные, отражающие информацию о контрагенте карточки договора, хранятся на вкладке «контрагент».

Полная сумма договора, а также сумма НДС в блоке RCM системы SAP указывается на вкладке «сумма договора».

На вкладке «контрировка» отображаются данные о центр финансовой ответственности (ЦФО), подразделение, контролирующее задолженность, банк, условие платежа договора.

Так же важным атрибутом карточки договора является вкладка «данные процедуры закупки». Она находится в узле связанных объектов, но без заполнения этих данных возможности перевести карточку договора в статус «на подписание» и далее в статус «в работе» возможности не будет, а следовательно, и проводить финансовые операции по договору тоже будет невозможно.

Для переноса карточки договора, необходимо заполнить шаблон с данными для переноса:

- вид договора;
- категория договора;

- дата заключения договора;
- дата начала действия договора;
- дата окончания платежей;
- дата окончания действия договора;
- номер договора контрагента;
- предмет договора;
- регистрационный номер в исторической системе;
- куратор договора;
- тип договора;
- идентификационный номер карточки договора в системе АСБУ;
- контрагент;
- центр финансовой ответственности (ЦФО);
- финансовая позиция.

После заполнения данного шаблона, с помощью инструмента загрузки, в систему SAP ERP загружаются карточки договоров в статусе «Черновик», поскольку куратор договора должен проверить основные данные карточки договора, прикрепить сканированный образ договора в формате PDF, создать бюджетный адрес для финансовых операций и после этого перевести карточку договора в статус «На подписание» и далее в статус «В работе». Только в данном статусе появляется возможность выполнять финансово-хозяйственные операции по данному договору.

Идентифицировать карточку договора и производить сверку исторической системы с системой SAP следует по полю «Регистрационный номер в исторической системе» и «ИД. Договора в системе АСБУ». Поскольку номер карточки договора в системе SAP представлен в формате XX-XX-XXXX (где первые два знака представляют собой принадлежность к балансовой единице, вторые два знака – это финансовый год, в котором заводится карточка договора, последние цифры представляют собой порядковый номер карточки договора в данном финансовом году в системе SAP).

Полный перечень атрибутов карточки, необходимый для переноса данных договора в систему SAP ERP представлен в приложении Б.

### **Перенос пользовательских настроек**

Таблицы конфигурации обновляются посредством выполнения одного из шагов IMG или непосредственно в одной из транзакций ведения ракурса. Эти стандартные процедуры предоставляют собой многофункциональное управление, которые включают в себя параметры проверки содержимого таблицы и автоматической регистрации изменений в запросе на перенос. Но на практике, частенько требуется ввод данных во все поля каждой новой записи при загрузке вручную при их применении.

При необходимости ввода большого количества новых записей в таблицу конфигурации, ввод возможен вручную или с применением автоматизации. Ручной ввод данных является долгим и трудозатратным, а также, при ручном вводе обновления таблицы конфигурации блокируются. Это происходит для того, чтобы только один пользователь мог вводить данные.

Для упрощения задачи ввода и автоматизации, рекомендуется использовать внешний инструмент или транзакцию, которую предлагают создатели системы SAP - SAP Legacy System Migration Workbench (LSMW). Но для использования этого инструмента необходимо сначала создать файл для ввода данных, а также записать его и протестировать скрипт, при этом сопоставить и сравнить каждое вводимое поле в файле и транзакции и только после этого можно запускать загрузку пакетным вводом. В данном случае вариант запуска возможен в фоновом и видимом режиме. Также при автоматическом процессе записи может возникнуть конфликт с методом загрузки, т.к. в

момент сохранения записи будет выведено диалоговое окно, в котором отражается выбор запроса на перенос.

### **Контроль за перенесенными данными по бухгалтерскому учету**

После перенесения данных бухгалтерского учета из исторической системы в систему SAP, необходимо осуществить контроль полученных данных.

Сверка данных производится по сальдо и оборотам на конкретных счетах, согласно ранее составленному маппингу с системой в определенном отчетном периоде.

Сверку данных необходимо выполнять по определенному набору контрировок, в зависимости от счета ими могут быть карточка договора, контрагент, ссылочные поля, такие как ИНН, КПП, ОКТМО, КБК, ссылочный ключ, как правило им является номер исходного документа (внешний номер счет-фактуры), МВЗ, МВП, заказ, код налога, финансовая позиция, ПФМ.

При внедрении рекомендуется сравнивать именно начальное сальдо, перенесенное с помощью шаблона в систему SAP ERP. Данный способ, на мой взгляд, наиболее целостно и четко показывает данные переноса и корректность их введения в систему SAP.

Для корректного переноса сальдо по бухгалтерскому учету, необходимо определить маппинг счетов ранее использованной системы и системы SAP ERP.

### **Выводы**

При принятии решения о переходе предприятия на систему SAP ERP, а также любую другую систему, руководство должно четко представлять себе тот объем работ, который предстоит проделать. Поскольку к каждому из пунктов необходимо относиться серьезно для успешного перехода в систему SAP ERP. Важным и трудозатратным является процесс написания алгоритма и программы по переносу данных предприятия, а также сверка перенесенных данных. Процесс усложняется и тем, что в различных системах одни и те же атрибуты представлены по-разному. Например, в некоторых системах, счета Главной книги представляются двухзначными числами, в то время как в системе SAP ERP счета Главной книги представляются из одиннадцати знаков. То же самое касается и карточки договоров, нумерация в разных системах ведется по-разному, что усложняет работу по переносу и выверке данных. Следовательно, особое внимание рекомендуется обращать на заполнение шаблонов миграции данных и их выверке, поскольку именно от этого зависит успех перехода в систему SAP ERP.

### **Библиографический список**

1. **Кречмер, Р.** SAPTEC: mySAP Technology / Изд. Лори, 2016.
2. **Рон, Хардман** Oracle Database PL/SQL. Рекомендации эксперта / Хардман Рон. - М.: ЛОРИ, 2017. – 8 - 9 с.
3. **Brück, U.** Praxishandbuch SAP-Controlling / Brück U. - SAP PRESS, 2015. - 586 с.
4. **Вейс, В.** BC405-BC470: SAP NetWeAver/ PArticipAnt HAndbook, 2015.
5. **Конноли, Т., Бегг, Л., Страчан, А.** Базы данных. Проектирование, реализация и сопровождение. Теория и практика. 3-е издание. Вильямс 2013. - Таблицы, картинки, 1436 с.
6. **Мамаев, Е.** MicrosoftSQLServer2000 - СПб.: БХВ-Петербург, 2018, 1280 с.

**ПРИМЕНЕНИЕ СТАТИСТИЧЕСКИХ МЕТОДОВ ОБНАРУЖЕНИЯ СЕТЕВЫХ АНОМАЛИЙ В АППАРАТНО-ПРОГРАММНОМ КОМПЛЕКСЕ «БЕЗОПАСНЫЙ ГОРОД»**

**О.Е. Виноградов,**

e-mail: [vinogradov-oe@yandex.ru](mailto:vinogradov-oe@yandex.ru),

**Е.А. Косолапов**

e-mail: [kosolapov00@yandex.ru](mailto:kosolapov00@yandex.ru)

Нижегородский государственный технический университет им. Р.Е. Алексеева

С целью реализации системного подхода к обеспечению общественной безопасности, правопорядка и безопасности среды обитания в условиях сохранения высокого уровня рисков природного и техногенного характера, а также сохраняющейся напряженной криминогенной обстановки в 2018 году в Нижегородской области создан аппаратно-программный комплекс «Безопасный город». Его реализация представляет собой комплексную информационную систему, обеспечивающую прогнозирование, мониторинг, предупреждение и ликвидацию возможных угроз. При решении задач управления такой сложной информационной системой востребованы такие технологии мониторинга, которые оказывали бы поддержку системным и сетевым администраторам в разрешении проблем сокращения среднего времени реакции на аномалии в объекте мониторинга. В представляемой системе в качестве управляющего объекта, в соответствии с теорией активных систем, используется открытое решение распределенного мониторинга – Zabbix. В статье рассмотрен опыт применения статистических методов обнаружения аномального поведения сетевого трафика.

**Ключевые слова:** АПК Безопасный город, информационная система, теория активных систем, обнаружение сетевых аномалий.

**APPLICATION OF STATISTICAL METHODS FOR DETECTING NETWORK ANOMALIES IN THE HARDWARE AND SOFTWARE COMPLEX "SAFE CITY"**

**O.E. Vinogradov, E.A. Kosolapov**

When solving problems of managing a complex information system such as the hardware and software complex "Safe City". Monitoring technologies are in demand that would support system and network administrators in solving problems of reducing the average response time to a malfunction. In this article, an open distributed monitoring solution, Zabbix, is used as a control object, in accordance with the theory of active systems. The experience of using statistical methods for detecting abnormal behavior of network traffic is considered.

**Keywords:** APK Safe City, information system, active systems theory, network anomaly detection.

## Введение

С целью реализации единого системного подхода к обеспечению общественной безопасности, правопорядка и безопасности среды обитания в условиях сохранения высокого уровня рисков природного и техногенного характера, а также сохраняющейся непростой криминогенной обстановкой и невысокой дисциплиной жителей, продолжающейся тенденции к урбанизации, руководством страны реализуются мероприятия, направленные на создание АПК «Безопасный город» [1].

АПК Безопасный город представляет собой комплексную информационную систему, обеспечивающую прогнозирование, мониторинг, предупреждение и ликвидацию возможных угроз, а также контроль устранения последствий чрезвычайных ситуаций и правонарушений с интеграцией под ее управлением действий информационно-управляющих подсистем дежурных, диспетчерских, муниципальных служб для их оперативного взаимодействия в интересах муниципального образования. Он 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]. При решении задач управления такой сложной информационной системой востребованы такие технологии мониторинга, которые оказывали бы поддержку системным и сетевым администраторам в разрешении проблем сокращения среднего времени реакции на неисправность.

В АПК Безопасный город в качестве управляющего объекта, в соответствии с теорией активных систем, используется открытое решение распределенного мониторинга – Zabbix. В статье рассмотрен опыт применения статистических методов обнаружения аномального поведения сетевого трафика.

## Теория активных систем

Вопрос о повышении эффективности управления такой информационной системой как АПК «Безопасный город» рассмотрен через призму теории активных систем (ТАС), которая является составной частью теории управления различными видами систем и изучает технологии их функционирования, обусловленные проявлениями активности участников системы. Основными методами исследования в ТАС являются математическое и имитационное моделирование. В течение тридцати лет развития теории активных систем было разработано, исследовано и внедрено множество эффективных организационных механизмов. Модели и методы ТАС находят применение при решении широкого круга задач управления в экономике и обществе – от управления технологическими процессами до принятия решения на уровне регионов и стран.

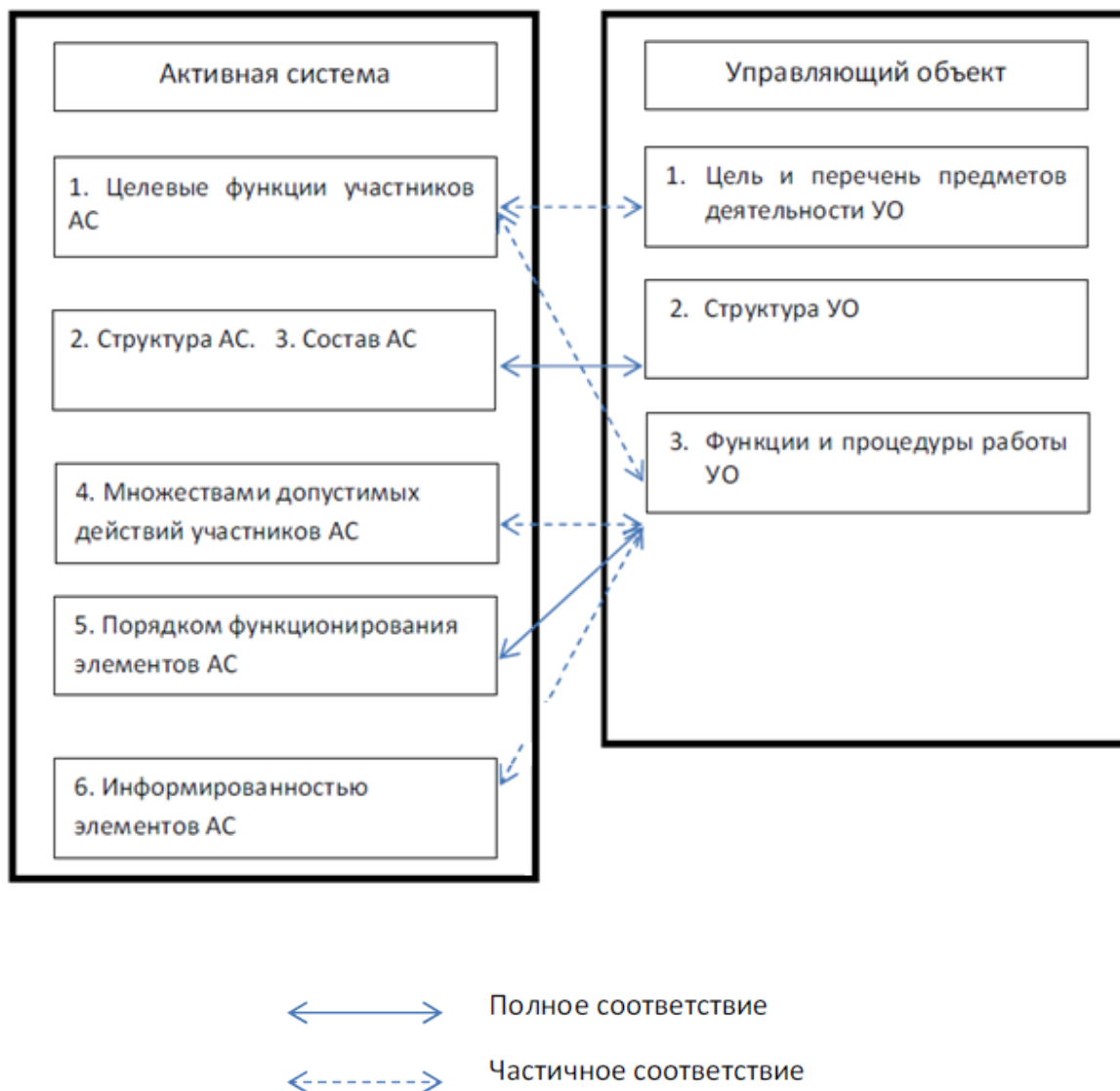
Рассмотрим задачу управления некоторой системой. Пусть состояние системы описывается переменной  $y \in A$ , принадлежащей допустимому множеству  $A$ . Состояние системы в некоторый момент времени зависит от управляющих воздействий со стороны управляющего объекта  $(UO) \quad u \in U : y = F(u)$ . Предположим, что на множестве  $U * A$  задан функционал  $\Phi(u, y)$ , определяющий эффективность функционирования системы. Величина  $K(u) = \Phi(u, F(u))$  называется эффективностью управления  $u \in U$ . Тогда задача управляющего органа заключается в выборе такого допустимого управления, которое максимизировало бы значение его эффективности при условии, что известна реакция системы  $F(u)$  на управляющие воздействия:

$$K(u) \rightarrow \max_{u \in U}$$

Следовательно, конкретизируется и модель системы  $F(u)$ , которая должна учитывать проявления активности управляемых субъектов. Управляемые субъекты стремятся к выбору таких своих состояний, которые являются наилучшими с точки зрения

их предпочтений при заданных управляющих воздействиях, а управляющие воздействия, в свою очередь, зависят от состояний управляемых субъектов. Одним из важнейших проявлений активности также является способность управляемых субъектов «предсказывать» (в рамках имеющейся информации) поведение управляющего органа – его реакцию на состояние системы и т.д.

Если управляющий орган имеет модель реальной активной системы, которая адекватно описывает ее поведение, то задача управления АС (задача синтеза оптимального управляющего воздействия) сводится к сформулированной выше – выбрать оптимальное управление  $u = \tilde{u}(y) \in U$ , то есть допустимое управление, максимизирующее эффективность.



**Рис. 1. Соотношение элементов систем**

Между отдельными элементами описания АС и УО можно установить определенные или частичные связи (рис. 1).

Перенос ТАС на современные системы мониторинга сложных информационных систем, модель активной системы будет задаваться перечислением следующих параметров АС:

- 1) целевыми функциями участников;
- 2) структурой;

- 3) составом входящих в АС элементов (участников);
- 4) множествами допустимых действий участников;
- 5) порядком функционирования элементов;
- 6) информированностью участников [3].

Управляющий объект принимает на себя ряд функций по управлению действиями участников и может быть описан следующими составляющими:

- 1) цель и перечень предметов деятельности;
- 2) структура;
- 3) функции и процедуры функционирования.

При постановке и решении задачи управления активной системой (АС), такой как АПК «Безопасный город» (управляемые субъекты обладают свойством активности, в том числе свободой выбора своего состояния) возникают следующие трудности. Кроме возможности выбора состояния, элементы активной системы обладают собственными интересами и предпочтениями, то есть осуществляют выбор состояния целенаправленно. Применим требования к функционированию активной информационной системы к работе современных систем мониторинга, выступающих в качестве управляющего объекта. В качестве управляющего объекта применяется система мониторинга – это совокупность технических средств, осуществляющих постоянное наблюдение и сбор информации в локальной вычислительной сети. В данной статье рассматривается применение системы мониторинга Zabbix. Позволяет собирать, хранить и отображать контролируемые показатели, контролировать нахождение параметров в допустимых пределах. Система мониторинга контролирует набор параметров коммутатора по протоколу SNMP.

Трафик с камер наружного наблюдения, рассредоточенных по городу, поступает через входящий сетевой интерфейс коммутатора на сервер обработки информации. Для контроля работоспособности системы камер производится анализ трафика от камер на входном интерфейсе коммутатора.

Кроме того, для трансляции видеопотока с камер на сервер используется кодек H264 — лицензируемый стандарт сжатия видео, предназначенный для достижения высокой степени сжатия видеопотока при сохранении высокого качества.



**Рис. 2. График данных с интерфейса**

Стандарт H.264 содержит ряд возможностей, позволяющих значительно повысить эффективность сжатия видео:



- Многокадровое предсказание;
- Использование сжатых ранее кадров в качестве опорных;
- Независимость порядка воспроизведения изображений и порядка опорных изображений;
- Независимость методов обработки изображений и возможности их использования для предсказания движения;
- Компенсация движения с переменным размером блока (от 16x16 до 4x4 пикселя) позволяет крайне точно выделять области движения;
- Векторы движения, выводящие за границы изображения;
- Шеститочечная фильтрация компонента яркости для полупиксельного предсказания с целью уменьшения зубчатости краев и, в конечном счёте, обеспечения большей чёткости изображения;
- Точность до четверти пикселя (Qpel) при компенсации движения обеспечивает очень высокую точность описания движущихся областей;
- Взвешенное предсказание, позволяющее использовать масштабирование и сдвиг после компенсации движения на величины, указанные кодером;
- Пространственное предсказание от краёв соседних блоков для I-кадров;
- Сжатие макроблоков без потерь;
- Функции преобразования: точное целочисленное преобразование пространственных блоков 4×4 и дополнительное преобразование Адамара;
- Квантование;
- Внутренний фильтр деблокинга в цикле кодирования, устраняющий артефакты блочности, часто возникающие при использовании основанных на DCT техниках сжатия изображений;
- Энтропийное кодирование квантованных коэффициентов трансформации;
- Функции устойчивости к ошибкам;
- Разбиение данных;
- Нумерация кадров [4].

В результате применения стандарта сжатия видео H.264 достигается высокая степень сжатия и как результат цикличное изменение сетевого трафика.

### **Статистический анализ объёмов трафика входного интерфейса АПК «Безопасный город»**

Проблемы при решении вопросов управления и защиты сетей вызваны сложностью логической и физической организации современных сетей. Администраторам в процессе эксплуатации компьютерных сетей приходится решать следующие задачи:

- Диагностика работы сети и подключенных к ней серверов, рабочих станций и соответствующего программного обеспечения;
- Обеспечение конфиденциальности, целостности и доступности.

При решении задач, связанных с диагностикой и защитой сетевых ресурсов, центральным вопросом является оперативное обнаружение состояний сети, приводящих к потере полной или частичной ее работоспособности, уничтожению, искажению, или утечке информации, являющихся следствием отказов, сбоев случайного характера или результатом получения злоумышленником несанкционированного доступа к сетевым ресурсам и других угроз информационной безопасности. Обнаружение таких состояний на ранней стадии позволит своевременно устранить их причину, а также предотвратит возможные катастрофические последствия.

Для решения вопросов диагностики сетей применяются систем управления, анализаторы сетевых протоколов, системы нагрузочного тестирования, системы сетевого мониторинга. Вопросы защиты сетей решаются с помощью межсетевых экранов, антивирусов и систем обнаружения атак, систем контроля целостности, криптографических средств защиты. Указанные средства диагностики и защиты

направлены на обнаружение известных и точно описанных типов воздействий, но малоэффективны в обнаружении модификации или новые типы воздействий.

В настоящее время актуальной задачей является поиск более эффективных методов выявления недопустимых аномалий в работе сети, являющихся следствием технических сбоев или несанкционированных воздействий. Главным требованием к этим методам является возможность обнаружения произвольных типов аномалий, в том числе новых, а также воздействий, распределенных во времени.

Современный подход заключается в поиске методов анализа, позволяющих выявлять аномальные состояния информационных ресурсов в виде отклонений от обычного состояния. Это позволяет обнаружить как известные, так и новые типы проблем [5].

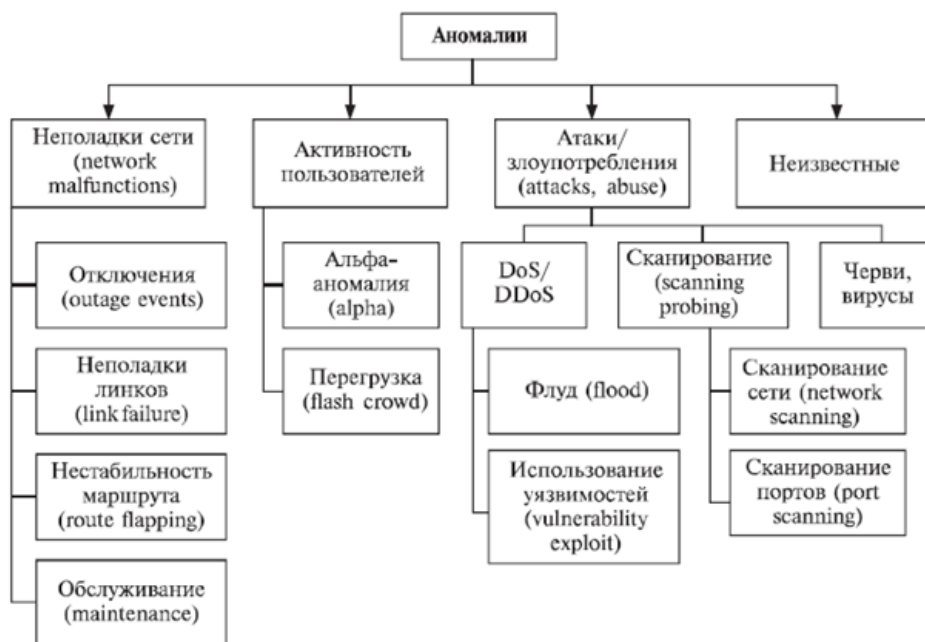


Рис. 3. Классификация аномалий в IP-сетях по типу

### Анализ методов обнаружения аномалий

При обнаружении аномалий необходимо решить следующие задачи:

- выбор оптимального множества параметров оценки;
- определение общего показателя аномальности с условием, что эта величина должна характеризовать общее состояние «аномальности».

Рассмотрим статистические методы обнаружения сетевых аномалий, практическое решение которых дает метод контрольных карт.

Контрольные карты Exponentially Weighted Moving Average (далее - EWMA) базируются на экспоненциальном сглаживании наблюдений.

Статистические показатели собираются на входном интерфейсе коммутатора за промежутки времени 48 часов с периодичностью 1 минута с использованием API интерфейса сервера Zabbix.

На основе полученных данных вычисляются допустимые верхняя и нижняя границы данного показателя:

$$\text{LowerLevel} = m - L \cdot d \cdot (\lambda/2 - \lambda)^{0.5},$$

$$\text{UpperLevel} = m + L \cdot d \cdot (\lambda/2 - \lambda)^{0.5},$$

где  $\lambda = 0.25$ ;  $L = 3$ ;  $m$  – математическое ожидание;  $d$  – среднеквадратическое отклонение.

После предварительных вычислений, в системе Zabbix создается «вычисляемый» элемент, соответствующий скользящему среднему значению с окном в 10 минут. Элемент имеет тип «вычисляемое» и формулу «avg("net.if.in [if In Octets.671088651]", 600)». Для

этого элемента устанавливаются триггеры, срабатывающие при пересечении границ (ранее вычисленных Lower Level, Upper Level).

В целях большей наглядности графики показателя, скользящего среднего и допустимых границ выведены на отдельную панель.



Рис. 4. График скользящего среднего данных с интерфейса

## Заключение

Таким образом, разработана методики обнаружения сетевых аномалий объема трафика с помощью универсальной системой мониторинга Zabbix с использованием метода контрольных карт Exponentially Weighted Moving Average для выявления неисправных или некорректно работающих узлов и оповещения ответственных лиц.

## Библиографический список

1. 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).
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](http://07.mchs.gov.ru/upload/site38/document_file/N0LDsp5FVR.pdf)(дата обращения: 26.05.2018).
3. **Бурков, В.** Теория активных систем: состояние и перспективы. / В. Бурков, Д. Новиков. - Москва 1999.
4. URL:<https://ru.wikipedia.org/wiki/H.264>.
5. **Шелухин, О.** Обнаружение вторжений в компьютерные сети. Сетевые аномалии. / О. Шелухин, Д.Сакалема, А. Филинова.

## ПОДХОД К ОПТИМИЗАЦИИ МНОЖЕСТВА ТЕХНОЛОГИЧЕСКИХ ОБЪЕКТОВ УПРАВЛЕНИЯ НА ОСНОВЕ МАШИННОГО ОБУЧЕНИЯ (НЕЧЕТКАЯ ЛОГИКА)

**Шамасна Хамза**

ORCID: 0000-0002-7637-2471 e-mail: t26032017@gmail.com

Нижегородский государственный технический университет им. Р.Е. Алексеева

**А.В. Семашко**

ORCID: 0000-0003-4687-9666 e-mail: semashko52@yandex.ru

Нижегородский государственный технический университет им. Р.Е. Алексеева

В работе рассматривается подход к использованию метода машинного обучения (с элементами нечеткой логики) для оценки целевых значений влажности и температуры подбора наиболее оптимальных целевых значений. Подход позволяет формализовать требования с применением аппарата нечеткой логики и определить частные целевые функции для поддержки решений.

**Ключевые слова:** Нечеткая логика, искусственный интеллект, машинное обучение, промышленный интернет вещей, мониторинг, большие данные, аналитика, принятие решений, интеллектуальное управление фермой, поддержка принятия решений, информационная система.

## MACHINE LEARNING (FUZZY LOGIC) APPROACH TO OPTIMIZING A SET OF TECHNOLOGICAL CONTROL OBJECTS

**H.A. Shamasna, A.V. Semashko**

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

This paper discusses an approach to using machine learning (Fuzzy Logic) to estimate the target values of humidity and temperature selection of the most optimal target values. The approach allows you to formalize the requirements using the apparatus of fuzzy logic and determine the private target functions to support decisions.

**Keywords:** Fuzzy logic, artificial intelligence, machine learning, industrial internet of things, monitoring, big data, analytics, decision making, intelligent farm management. decision support, information system

## Введение

В последнее десятилетие информационные системы в сочетании с Интернетом, облачными вычислениями, мобильными устройствами и технологиями «Интернет вещей» привели к появлению огромных объемов данных, обычно называемых большими данными. Они включают структурированные, слабоструктурированные и неструктурированные данные реального времени. Информатика продвинулась вперед в хранении и обработке больших объемов разнообразных наборов данных с использованием статистических методов решений в режиме реального времени для улучшения бизнес-результатов. Революция больших данных сильнее аналитики, которая использовалась в прошлом. Заинтересованные деловые компании и ученые разработали уникальные способы извлечения выгоды из больших данных [1].

Одна из областей, в которой весьма эффективным представляется внедрение сервисов «Интернета вещей» (IoT) и 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 птицеферм 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.

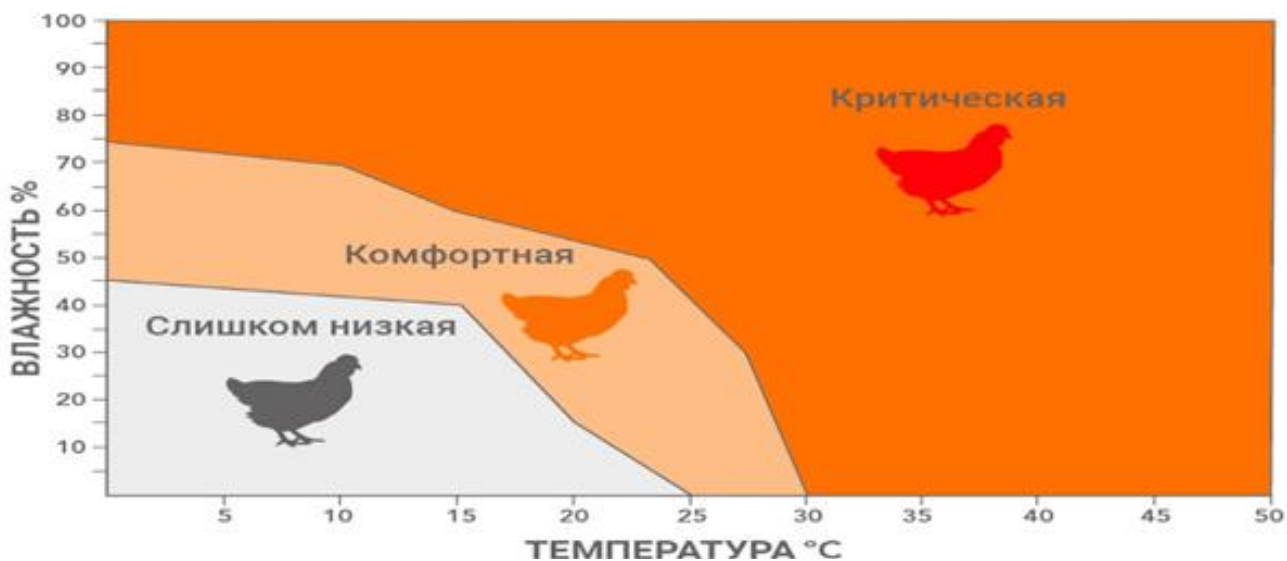
Крупные хозяйства приступают к внедрению комплексных платформ управления процессами на базе IoT — интеллектуальных сетей датчиков, видеокамер, беспилотных аппаратов и других устройств, искусственного интеллекта (ИИ), облачных технологий и робототехники. В частности, платформа Episcopy, разработанная компанией Wave Access на базе технологий IoT и ИИ, помогает анализировать параметры эффективности 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 и рассчитывают показатели (метрики) 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» [2]. Например, от качественно приготовленного комбикорма зависит более 60% рентабельности всего производства [3]. 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 позволит выработать рекомендации по управлению режимами функционирования для множества ферм и птицефабрик, объединенных сетью промышленного IoT.

Применение технологий промышленного интернета вещей позволяет перейти от независимого управления отдельными типовыми объектами управления (ТОУ) к управлению совокупностью ТОУ. Обработка статистического ансамбля реализаций, формируемых от

совокупности ТОУ, позволяет значительно снизить уровень неопределенности и получить потенциальный выигрыш по точности управления, таким образом, предлагается использовать технологии промышленного IoT для объединения множества территориально распределенных ТОУ в одну систему, что позволит централизованно обрабатывать данные мониторинга и кардинально повысить качество управления [4].

Система управления микроклиматом предназначена для автоматического поддержания параметров микроклимата (температура, влажность, загазованность, и т.д.) в заданных пределах в течение всего цикла выращивания птицы. Целевые значения влажности и температуры могут быть найдены, как множество значений с возрастом (по мере роста птицы). Известно, что существует так называемая «термо-нейтральная зона», представляющая собой такой диапазон температур окружающей среды, в пределах которого птица чувствует себя комфортно и не тратит свою энергию на терморегуляцию тела. Комфортные условия содержания птицы определяется областью значений влажности и температуры воздуха (рис.1) [5]. При различной влажности воздуха птица по-разному ощущает одну и ту же температуру. Область значений комфортного содержания, определена приближенно эмпирическим путем и может различаться в зависимости от вида выращиваемых птиц и других параметров условий содержания, в том числе и неконтролируемых. Управление целевыми значениями (влажности и температуры) в множества ТОУ является ключевой задачей, поскольку каждый ТОУ имеет свои собственные характеристики (рост, внешне среда, занятости, и.т.д.) с учётом динамики и изменения целевых значений каждого объекта индивидуально со временем в зависимости от собственных характеристик [5]. Применение машинного обучения и ИИ позволяет плавно менять целевые значения для управляемых параметров, используя результаты мониторинга для множества птичников, которые собираются благодаря IoT.



**Рис.1. Соотношение влажности и температура, и их влияние на кур [5]**

Целевые значения влажности и температуры могут быть найдены, как множество значений, равноудаленных от границ. При этом предпочтительная область пар значений (рабочий отрезок) находится в центре – в диапазоне 16...20 градусов. С возрастом (по мере роста птицы) целевые значения температуры и влажности изменяются. Целевые значения представлены в табл. 1 [6].

Таблица 1

## Целевые значения влажности и температуры [6].

Вес (г)	30%	40%	50%	60%	70%	80%
42	33	32.2	32	29.5	29	27
175	33	32.2	32	29.5	29	29
486	30	31	31	29	28	26.5
931	28	28	27.5	26.5	26	25
1467	26	25	25	24	23.5	22.5
2049	23	23	22.5	22	21	20.5
2634	20	20	19.5	18.5	17.5	16
3177	18	17.5	17	16	15	14
4064	14	13.5	13	12	11	10

Использование технологий, основанных на ИИ, для автоматического и точного сбора данных в режиме реального времени способствует углубленному анализу, который может позволить немедленно принять меры по оптимизации производственных показателей фермы. ИИ может извлекать уроки из этих данных как под наблюдением, так и без него. При контролируемом обучении ИИ обучается на уже существующих наборах данных, которые направлены только на определенную управляемую работу.

Один из основных сдерживающих факторов развития животноводства является тепловой стресс летом из-за высокой температуры, влажности и углекислого газа, на практике большая часть фермеров использует традиционные стратегии контроля с низкой эффективностью, что является следствием игнорирования сильных взаимодействий между бойлерами и окружающей средой (внешним климатом) в том смысле им приходится платить дополнительную стоимость за потребление энергии, используемой исполнительными механизмами, чего во многих случаях можно избежать, таким образом это, как правило, неэффективно, вредно для здоровья и может привести к значительным экономическим потерям [6].

В настоящее время в высокотехнологичных хозяйствах практикуется использование широкого спектра датчиков для измерения массы птицы, температуры, расхода корма и воды, влажности, уровня аммиака, CO<sub>2</sub> и многих других параметров. Контролируемые компьютером машины и робототехника могли бы значительно сократить взаимодействие человека с сельскохозяйственными птицами, уменьшить источник инфекции, а также повысить эффективность работы по сравнению с ручным трудом.

Использование искусственного интеллекта может снизить частоту ошибок до незначительной и работать 24×7, таким образом, он может повысить эффективность сельского хозяйства, что приведет к максимизации прибыли фермы.

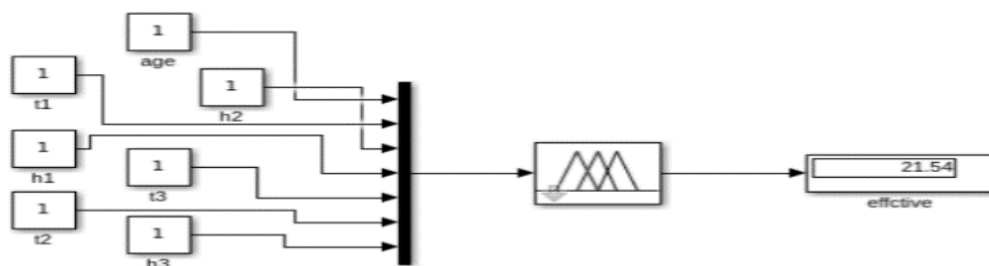
Искусственный интеллект в ближайшем будущем может превратить нынешнее лицо традиционного промышленного земледелия в умное птицеводство или птицеводство с использованием искусственного интеллекта. Нечеткая логика обладает большим потенциалом для исследователей, и она была разработана за последние два десятилетия, является методом машинного обучения, используя человеческие знания и опыт для обработки информации в базах правил и получения результата, что обеспечивает поддержку принятия решений. Классическая логика по своей сути игнорирует проблему неопределенности, поскольку все высказывания и рассуждения в формальных логических системах могут иметь только значение "истина" (И, 1), Нечеткая логика, которая предназначена для формализации человеческих способностей к неточным или приближенным рассуждениям, позволяет более адекватно описывать ситуации с неопределенностью.

На основе нечеткой логики предполагается оптимизация, суть которой в том, чтобы обеспечить возможность машинного обучения для плавного изменения целевых значений управляемых параметров, используя результаты мониторинга в реальном времени для

множества птичников, которые собираются благодаря IoT.

В рамках разработки интеллектуальной системы консультационной поддержки оптимальных решений предполагается, что на основе запроса эксперта определяются Целевые значения, информация о которых в дальнейшем использоваться в процессе интерактивной визуализации. Под интерактивной визуализацией понимается организация обратной связи таким образом, что, предлагаемые системой управления, будут актуальны для эксперта в рамках всего цикла выращивания птицы и будут оптимизировать определенные целевые функции.

Методология проектирования, методы нечеткой логики реализованы на платформе (MATLAB&SIMULINK) (Рис.2.). Первым шагом в применении нечеткой логики является формулировка проблемы и определение соответствующих параметров, Вход 1 – это возраст (age). Она делится на 3 треугольные функции принадлежности и первую, вторую и третью недели (week1, week2, week3), изменяющимся для каждого возраста. Целевые значения влажности и температуры, Вход 2 — это температура (Tempreture). Она делится на 5 треугольных функций принадлежности и 1 трапециевидную функцию принадлежности, которые называются (очень холодной, холодной, комфортной, горячей и очень горячей (very cold, cold, hot and very hot). Входы 3 это влажности (Humedity). Она делится на 5 треугольных функций принадлежности и 1 трапециевидную функцию принадлежности, которые называются очень низкий, низкий, комфортный, высокий, очень высокий, (very low, low, comfort , high , very high) , важно сказать что для каждого возраста меняются целевые значения влажности и температуры, а затем способ формирования функций принадлежности. Эта система будет использовать механизм вывода Мамдани. Между Входом 1, Входом 2 и Входом 3 будет применено 72 правила [7,8 9,10,11]. В качестве выходного показателя системы были использованы показатели эффективности, которые характеризуют качество реализации проекта. Они делится на три названные ниже: низкий, средний, высокий (bad, good, verygood). Функции принадлежности – Рис. 3. Матрица правил – Рис. 4.



**Рис.2. Проектирование нечеткой логики на платформе MATLAB&SIMULINK**



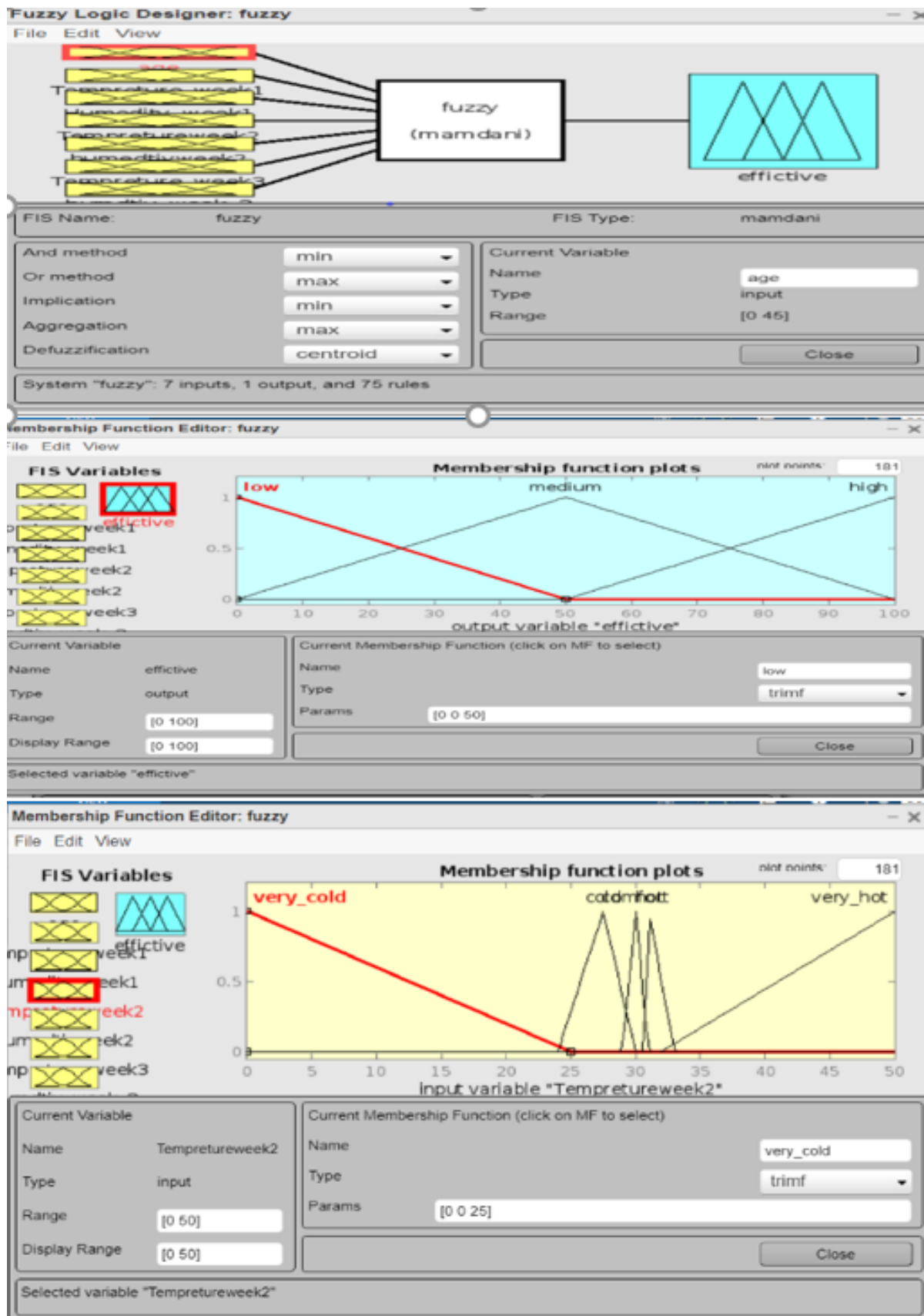


Рис. 3. Входные и выходные функции принадлежности

	холодно,	низкий	низкий	средний	низкий	низкий
комфортный		средний	средний	высокий	средний	низкий
жарко		средний	средний	средний	низкий	низкий
Очень жарко		низкий	низкий	низкий	низкий	низкий

Рис.4. Матрица правил

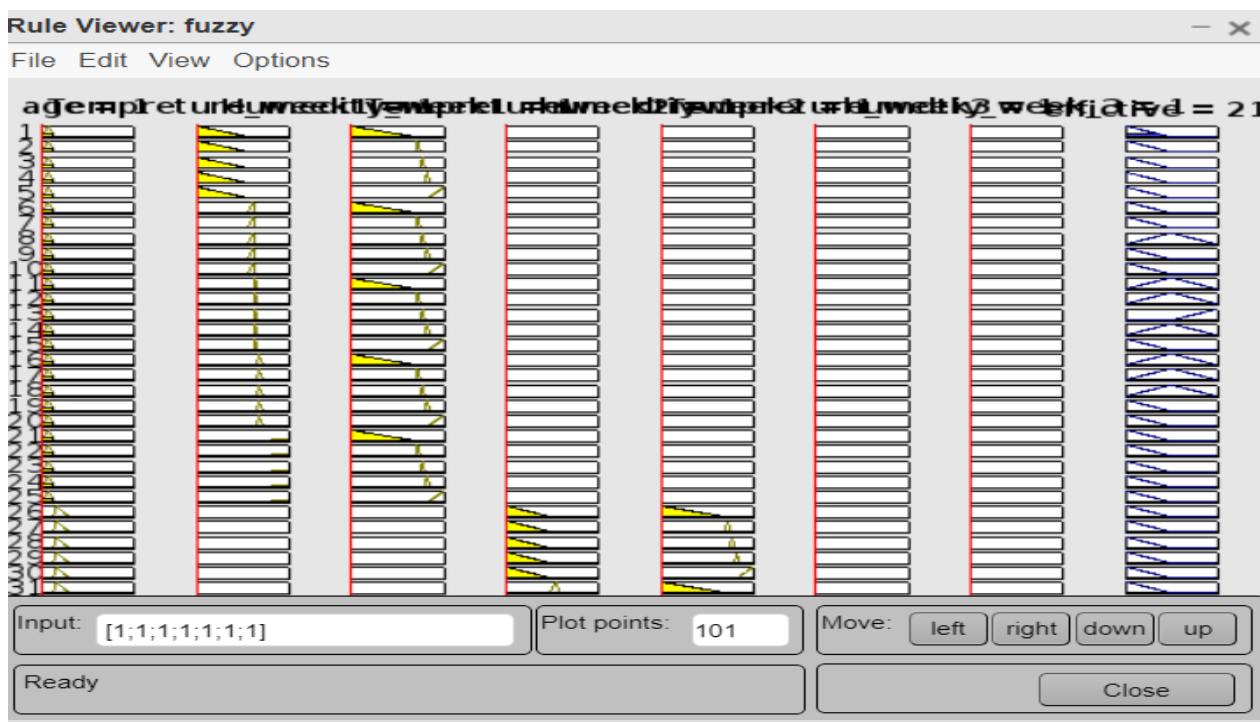


Рис.5. Правила и результат работы системы в программе просмотра правил MATLAB

На последнем этапе определяется метод преобразования нечетких чисел, генерируемых моделью нечеткой логики, в классические числа [12]. Важным шагом в нечетком моделировании и принятии нечетких решений является процесс дефаззификации, который определяет наилучшее нечеткое значение производительности. Доступно несколько методов такой дефаззификации, включая среднее значение максимумов (MOM), [13]. В качестве метода дефаззификации использовался метод MOM, который вычисляет среднее значение максимумов в нечетких выходных данных.

Практика показывает, что эксперты выявляют оптимальные условия неточно. Соответственно, при обработке больших данных разных исследований экспертов, возможно получить более эффективные решения целевых значений влажности и температуры, приведенные в Таблице 2.

**Таблица 2**

**Оптимальные целевые значения влажности и температуры**

Вес (г)	30%	40%	50%	60%	70%	80%
42	32.04	31.75	31.42	30.04	29.26	28.20
175	32.37	32.08	31.76	30.37	29.59	28.54
486	30.84	30.55	30.22	28.84	28.06	27.00
931	28.42	28.13	27.81	26.42	25.64	24.59
1467	25.92	25.63	25.31	23.92	23.14	22.09
2049	23.59	23.30	22.97	21.59	20.81	19.75
2634	20.17	19.88	19.56	18.17	17.39	16.34
3177	17.84	17.55	17.22	15.84	15.06	14.00
4064	13.84	13.55	13.22	11.84	11.06	10.00

По мнению исследователей нечеткая логика может успешно использоваться в птицеводстве. Способность распознавать и определять состояния системы привела к ее быстрому использованию в областях применения. Подходы ИИ, такие, как нечеткая логика, могут быть инициированы опытом и интуицией эксперта, поскольку в них включена информация, извлечённая в результате интеллектуального анализа большого массива экспертных данных. Благодаря этой особенности нечеткая логика может быть альтернативным методом для исследований, связанных с зоотехнической наукой.

Ожидается, что предложенный подход использования машинного обучения, основанный на использовании нечеткой логики для обработки данных мониторинга от множества ТОУ, позволит оптимизировать и плавно менять целевые значения для управляемых параметров, используя результаты мониторинга для множества птичников, которые собираются, благодаря IoT.

**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. **Третьяков, С.А., Шержуков, Е.Л.** Умная птицеферма: взгляд в ближайшее будущее // Автоматизация в промышленности. 2009. № 1. С. 41 – 44.
3. Диспетчеризация в сельском хозяйстве [Электронный ресурс] URL:<https://smoking.com/a/121476/dispetcharizatsiya-v-selskom-khozyaystve/>
4. **Шамасна, Х.А.** Архитектура многоуровневой сетевой системы управления умными фермами// Информационные системы и технологии (ИСТ-2021): Сборник материалов XXVII Междунар. науч.-техн. конф. – Н. Новгород: НГТУ им. Р.Е. Алексеева, 2021,
5. <https://sagrada.biz/press-centr/stati/temperature-v-ptichnike/>

6. <https://sagrada.biz/press-centr/stati/mikroklimat-na-ptitsefabrike/>
7. **Wang, L:** A course in fuzzy systems and control. 108-115, Prentice Hall, Upper Saddle New Jersey, 1997.
8. **Elmas, Ç:** Fuzzy logic controllers. 1st ed., 35-40, Seçkin Press, Ankara, 2003.
9. Ross TJ: Fuzzy logic with engineering applications, 2nd ed., 134-137, John Wiley&Sons Ltd, Chichester, 2004.
10. **Akkaptan, A.** Hayvancılıkta bulanık mantık tabanlı karar destek sistemi. Yüksek Lisans Tezi, Ege Üniv. Fen Bil. Enst., 2012.
11. **Sivanandam, S.N., Sumathi, S., Deepa, S.N.** Introduction to Fuzzy logic using MATLAB. Springer, Berlin, 430, 2007.
12. **Şen, Z.** Principles of fuzzy logic and modeling. 2nd ed., Bilge Kültür Sanat Press, 2009. 2
13. **Liu, H.C., You J.X., You X.Y., Shan M.M.** A novel approach for failure mode and effects analysis using combination weighting and fuzzy VIKOR method. Appl Soft Comput, 28, 579-588, 2015. DOI: 10.1016/j. asoc.2014.11.036

**ПРОБЛЕМА РАЗРАБОТКИ ТЕОРЕТИЧЕСКИХ И ОРГАНИЗАЦИОННО-МЕТОДИЧЕСКИХ ОСНОВ АДАПТИВНОГО УПРАВЛЕНИЯ ЗАЩИЩЕННОСТЬЮ ОБЪЕКТОВ ИНФОРМАЦИОННОЙ ИНФРАСТРУКТУРЫ АВТОМАТИЗИРОВАННЫХ СИСТЕМ ОТ ВИРУСНЫХ АТАК**

**Р.А. Хворов**

ORCID: 0000-0002-5718-6227 e-mail: khvoroff@rambler.ru

Военный учебно-научный центр Военно-воздушных сил «Военно-воздушная академия им. проф. Н. Е. Жуковского и Ю. А. Гагарина» (г. Воронеж)

Предпосылки решения проблемы разработки теоретических и организационно-методических основ адаптивного управления защищенностью объектов информационной инфраструктуры автоматизированных систем (АС) от вирусных атак позволяют сформулировать постановку и порядок ее решения.

Требуется синтезировать множество характеристик защищенности объектов информационной инфраструктуры АС от вирусных атак, адекватность оценки управления защищенностью при котором будет превосходить адекватность оценки существующими методами управления.

Сформулированную проблему целесообразно решать в виде следующей последовательности задач:

- синтеза системы характеристик, отражающих возможности по обеспечению защищенности объектов информационной инфраструктуры АС от вирусных атак при заданных характеристиках, отражающих возможности нарушителя и характеристиках, отражающих возможности антивирусных механизмов;
- разработки системы математических моделей оценки исследуемых характеристик;
- обоснования показателя качества управления защищенностью объекта информационной инфраструктуры АС от вирусных атак;
- исследования влияния адаптивного управления защищенностью объекта информационной инфраструктуры АС от вирусных атак на качество их функционирования.

**Ключевые слова:** адаптивное управление, вирусные атаки, защищенность, автоматизированные системы.

**THE PROBLEM OF DEVELOPING THEORETICAL, ORGANIZATIONAL AND METHODOLOGICAL FOUNDATIONS OF THE REACTION OF MANAGING SECURITY OF OBJECTS DISCLOSURE CONFIDENTIAL SECURITY SYSTEMS FROM VIRUS ATTACKS**

**R.A.Khvorov**

Air force academy N.E. Zhukovsky and Yu.A. Gagarin»

**Purpose:** The prerequisites for solving the problem of developing theoretical, organizational and methodological foundations for adaptive control of the security of information infrastructure objects of automated systems (AS) from virus attacks allow us to formulate the formulation and procedure for its solution.

**Design/methodology/approach:** It is required to synthesize a set of characteristics of the security of AS information infrastructure objects from virus attacks, the adequacy of the assessment of security management in which will exceed the adequacy of the assessment by existing management methods.

**Research limitations/implications:** It is advisable to solve the formulated problem in the form of the following sequence of tasks:

- synthesis of a system of characteristics that reflect the ability to ensure the security of NPP information infrastructure objects from virus attacks with given characteristics that reflect the capabilities of the intruder and characteristics that reflect the capabilities of anti-virus mechanisms;
- development of a system of mathematical models for assessing the characteristics under study;
- substantiation of the quality indicator for managing the security of the NPP information infrastructure object from virus attacks;
- studies of the impact of adaptive control of the security of the NPP information infrastructure object from virus attacks on the quality of their functioning.

**Originality/value:** Improving the effectiveness of protection of information infrastructure objects of automated systems in the conditions of virus attacks.

**Keywords :** adaptive control, virus attacks, security, automated systems.

## **Введение**

Анализируя в целом существующий уровень проработки математических методов как в проблематике информационной безопасности в целом, так и в вопросах адаптивного управления защищенностью объектов информационной инфраструктуры автоматизированных систем (АС) от угроз вирусных атак, можно выявить два весьма серьезных обстоятельства, не позволяющих характеризовать разработанные на настоящий момент математический аппарат как адекватный.

### **Недостатки существующих моделей**

Первое обстоятельство связано с неоправданно широким применением эвристического подхода к оценке возможностей нарушителя по реализации угроз вирусных атак. Характерным примером здесь может служить методика ФСТЭК России [1] по оценке указанных возможностей, реализующая экспертный подход в вопросах вероятностного представления субъектно-объектных отношений в исследуемых процессах. Примененный здесь эмпирический подход к переходу от лингвистического к количественному представлению вероятностных характеристик угроз вирусных атак способствует доминированию субъективного влияния экспертных оценок, как на сам процесс исследования, так и на его результаты, что является предпосылкой низкой адекватности используемого методического аппарата. Кроме того, данный методический аппарат не учитывает те случайные состояния исследуемых процессов, которые характеризуют их динамику.

Второе обстоятельство связано с тем, что в тех моделях, которые все же позволяют оценить характеристики исследуемых процессов в их динамике [2], учитывается лишь продолжительность этих процессов, но не учитываются случайные состояния, связанные с динамикой возникновения угроз и их обнаружением. Как и в предыдущем случае, это обстоятельство является предпосылкой низкой адекватности этих моделей.

Преодоление указанных недостатков возможно лишь в том случае, когда формальное описание исследуемых процессов позволяет математически представить все условия, характерные для динамики вирусных атак и реагирования на них антивирусными механизмами.

Характеризуя, в этой связи, исследовательский потенциал методологии информационной безопасности в целом и методического аппарата управления защитой информации в частности, можно констатировать, что исследования в данной области позволили дать всестороннюю характеристику как угроз вирусных атак, так и способов защиты информации от такого рода угроз. Вместе с тем, методические подходы к исследованию качества управления защищенностью информации от угроз вирусных атак [3], не обладают той степенью системности, которая позволила бы учесть особенности динамики вирусных атак и механизмов антивирусной защиты в рамках однотипного формализованного представления исследуемых процессов.

Реализованные к настоящему времени подходы к систематизации проявлений эффекта защищенности информации от вирусных атак носят субъективный характер и в крайне ограниченном виде отражают системные проявления субъектно-объектных отношений в процессах, характерных для данной предметной области, включая формальную интерпретацию динамики вирусных атак и механизмов антивирусной защиты.

Возрастающие требования к противодействию вирусным атакам на информационные ресурсы АС в целом и требования к обоснованности мер антивирусной защиты в условиях современных тенденций совершенствования управления защищенностью информации, в частности, вопросы оценки защищенности являются актуальными и нуждаются в проработке как в методическом, так и в прикладном плане.

Как показывает анализ состояния проблемы, одним из наиболее перспективных путей ее решения является синтез системы характеристик процессов обеспечения

антивирусной защиты в АС в рамках целевой функции управления защищенностью от такого рода угроз.

### Постановка задачи

Изложенные выше предпосылки решения проблемы разработки теоретических и организационно-методических основ адаптивного управления защищенностью объектов информационной инфраструктуры АС от вирусных атак позволяют сформулировать постановку и порядок ее решения.

Формальная постановка проблемы состоит в следующем.

#### Заданы:

- множество ( $V$ ) функций, выполняемых в процессе реализации угроз вирусных атак на информационные ресурсы объекта информационной инфраструктуры;
- множество ( $R$ ) функций, выполняемых в процессе реагирования на такого рода угрозы безопасности информации;
- показатель ( $D$ ) защищенности объекта информационной инфраструктуры АС от вирусных атак;
- показатель  $C(D)$  качества управления защищенностью объекта информационной инфраструктуры АС от вирусных атак
- функция адекватности ( $A$ ) оценки качества управления защищенностью объекта информационной инфраструктуры АС от вирусных атак.

**Требуется:** синтезировать множество ( $X$ ) характеристик защищенности объекта информационной инфраструктуры АС от вирусных атак, адекватность оценки управления защищенностью при котором будет превосходить адекватность оценки существующими методами управления.

С целью формализации условий решения данной проблемы определим следующие множества:

- $X_{(V)}$  – множество характеристик, отражающих возможности нарушителя по реализации угроз вирусных атак на информационные ресурсы объекта информационной инфраструктуры;
- $X_{(R)}$  – множество характеристик, отражающих возможности антивирусных механизмов по реагированию на вирусные атаки;
- формируемое на основе множеств  $X_{(V)}$  и  $X_{(R)}$  множество  $X_{(D)}$  характеристик, отражающих возможности по обеспечению защищенности объекта информационной инфраструктуры АС от вирусных атак.

Кроме того определим множество вариантов  $w_j \in W, j = 1, 2, \dots, |W|$ , существующих методических подходов к оценке качества управления защищенностью информации в классе систем, к которым относятся объекты информационной инфраструктуры АС.

Тогда условием повышения адекватности оценки качества управления защищенностью объекта информационной инфраструктуры АС (при заданном множестве ( $V$ ) функций, выполняемых в процессе реализации угроз вирусных атак на информационные ресурсы объекта информационной инфраструктуры и множестве ( $R$ ) функций, выполняемых в процессе реагирования на такого рода угрозы) является условие:

$$A[C(W_j)] < A[C(X)], j = 1, 2, \dots, |W|, \\ |X| = Const.$$

## **Вывод**

Таким образом, сформулированную проблему целесообразно решать в виде следующей последовательности задач:

- синтез системы характеристик, отражающих возможности по обеспечению защищенности объекта информационной инфраструктуры АС от вирусных атак при заданных характеристиках, отражающих возможности нарушителя по реализации угроз вирусных атак на информационные ресурсы объекта информационной инфраструктуры и характеристиках, отражающих возможности антивирусных механизмов по реагированию на такого рода угроз;
- разработка системы математических моделей оценки исследуемых характеристик;
- обоснование показателя качества управления защищенностью объекта информационной инфраструктуры АСУ СН от вирусных атак;
- исследование влияния адаптивного управления защищенностью объекта информационной инфраструктуры АСУ СН от вирусных атак на качество управления за счет повышения адекватности его оценки.

## **Библиографический список**

1. Методика оценки угроз безопасности информации. Информационное сообщение об утверждении Методики оценки угроз безопасности информации от 15 февраля 2021 г. № 240/22/690. М.: ФСТЭК России, 2021 г. 83 с.
2. **Мещерякова, Т.В.** Аналитические модели показателей состояния защищенности информации в центрах обработки данных органов внутренних дел / Т.В.Мещерякова, М.Е. Фирюлин, Р.А.Хворов // Вестник Воронежского института МВД России. Воронеж: ВИ МВД России. 2015. № 3. С. 104-113.
3. **Скрыль, С.В.** Оценка защищенности информации от вирусных атак: существующий и перспективный методический аппарат. Промышленные АСУ и контроллеры / С.В.Скрыль, Т.В.Мещерякова, Д.А.Голубков, В.И.Арутюнова. М: Научтехлитиздат, 2018. №9. С. 51–62.



## ОПТИМИЗАЦИЯ ТРАНСПОРТНОЙ ЗАДАЧИ С НЕЧЕТКИМИ ДАННЫМИ

О.П. Тимофеева

ORCID: 0000-0002-1935-7697 e-mail: [optimofeeva@mail.ru](mailto:optimofeeva@mail.ru)

А.Н. Санников

ORCID: 0000-0002-8418-506X e-mail: [lexsannikov@yandex.ru](mailto:lexsannikov@yandex.ru)

В.А. Веселов

ORCID: 0000-0003-2498-246X e-mail: [idrди2@yandex.ru](mailto:idrди2@yandex.ru)

А.Ю. Карпычева

ORCID: 0000-0001-9236-5207 e-mail: [angelina00548@mail.ru](mailto:angelina00548@mail.ru)

Л.М. Кочеганова

ORCID: 0000-0003-1205-6637 e-mail: [lili-rashel@yandex.ru](mailto:lili-rashel@yandex.ru)

Нижегородский государственный технический университет им. Р.Е. Алексеева

В условиях становления цифровой экономики широкое применение получили задачи оптимизации логистических сетей поставок продукции, которые основаны на теории и методах решения транспортных задач. Одним из критериев оптимальности плана поставок продукции является минимизация издержек по ее транспортировке с целью уменьшения себестоимости товара на экономическом рынке. Большинство из существующих математических методов, оказываются неэффективными при решении управленческих задач в условиях недостаточной информации об исследуемых процессах и явлениях. Для определения оптимального производства и распределения продукции, согласно потребительскому спросу, используются элементы нечеткой логики. В статье приводятся постановка транспортной задачи с нечетко заданными потребностями, описание метода приближенного решения поставленной задачи, результаты тестирования на различных начальных условиях и сравнительный анализ результатов между классической транспортной задачей и задачей с нечеткими входными данными.

**Ключевые слова:** транспортная задача, нечеткая логика, нечеткие данные, ограниченность ресурсов.

## OPTIMIZATION OF A TRANSPORT TASK WITH FUZZY DATA

O.P. Timofeeva, A.N. Sannikov, E. Kasko, L.M. Kocheganova

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

**Purpose:** The purpose of this article is to formulate a transport problem in conditions of uncertainty of consumer demand and to simulate the operation of the algorithm under various initial conditions.

**Design/methodology/approach:** The concept of an algorithm for the operation of a logistics network for modeling a transport problem with fuzzy initial data to minimize financial costs and meet consumer demand is proposed.

**Findings:** As a result, a comparative analysis was made between the classical transport problem and using fuzzy data.

**Research limitations/implications:** The results obtained in the course of the study can be used by logistics companies to optimize the work on the supply of products.

**Originality/value:** The novelty of the work lies in the use of fuzzy logic elements to solve a transport problem with changing consumer demand.

**Keywords:** transport problem, fuzzy logic, fuzzy data, limited resources.

## Введение

Информационные технологии в наши дни играют огромную роль в развитии современного информационного общества и способствуют решению многих задач. Современные вычислительные мощности систем позволяют решать сложные задачи за считанные секунды и все чаще можно наблюдать, что уже повседневные задачи решаются с помощью различных вычислительных систем.[1]

В теории линейного программирования транспортные задачи играют особую роль и имеют высокую практическую значимость в связи с тем, что позволяют сократить транспортные издержки предприятия. Стремление производителей продукции к оптимизации любых затрат во все времена было актуальным, сейчас решение задачи минимизации издержек находит еще больший спрос, т.к. напрямую влияет на повышение конкурентоспособности предприятий - ведь тогда издержки покрываются меньшей частью прибыли и позволяют снизить себестоимость продукции на рынке. [2]

В области исследования операций в части решения транспортных задач к настоящему времени достигнуты весьма значительные результаты. Тем не менее, большинство из существующих математических методов оказывается эффективными при решении управленческих задач лишь в условиях наличия достаточной информации об исследуемых процессах и явлениях.[3] Однако, на практике зачастую приходится иметь дело с информацией, которая не является полностью определенной, а носит приближенный, неточный характер в условиях временных ограничений.[4] В таком случае на помощь приходит теория нечетких множеств, методы которой являются общепринятыми инструментами для решения самого широкого спектра задач принятия решений в условиях субъективности информации о внешней среде и недостаточной априорной определенности [5].

### Постановка задачи логистики с нечетко заданными потребностями

Рассмотрим задачу логистики, в которой потребности являются нечеткими и задаются с помощью нечетких величин. Будем предполагать, что снабжение  $m$  клиентов однородными материальными средствами производится с  $n$  заводов, их потребности  $b_1^*, b_2^*, \dots, b_m^*$  являются нечеткими числами и равны примерно  $b_1, b_2, \dots, b_m$ . [3] Нечеткие величины  $b_1^*, b_2^*, \dots, b_m^*$  задаются на универсальном множестве  $U = [0, b']$ , где  $b'$  — достаточно большое число, такое что  $b' \geq b_j$  при любом  $j=1, \dots, m$ . Клиент с номером  $j$  готов согласиться на некоторое уменьшение своих потребностей по сравнению с величиной  $b_j$ ; при этом  $s_j$  — максимальная величина недопоставки, на которую он может согласиться. [6] Функции принадлежности определенных таким образом нечетких потребностей обозначим  $\mu_{b_j^*}(x)$ ,  $j=1, 2, \dots, m$ ; тогда для объема поставок  $x \in [0, b']$ :

$$\mu_{b_j^*}(x) = \begin{cases} 0, & \text{если } 0 \leq x \leq b_j - s_j, \\ \frac{x - (b_j - s_j)}{s_j}, & \text{если } b_j - s_j \leq x \leq b_j, \\ 1, & \text{если } b_j \leq x \leq b'. \end{cases}$$

Чем больше величина уменьшения потребности  $s_j$ , тем меньше значение функции принадлежности, которая представляет собой степень уверенности в том, что клиент удовлетворен объемом поставки.[7]

Формализованные таким образом (1) нечеткие потребности - трапецеидальные нечеткие числа:

$$\langle b_j - s_j, b_j, b' \rangle, j=1, 2, \dots, m.$$

Запасы материальных средств на заводах (складах) равны  $a_1, a_2, \dots, a_n$ . Будем полагать, что для удовлетворения потребностей в минимальных объемах запасов достаточно, если это

не так, то будем считать, что задача не имеет допустимых решений. Но, даже если запасов хватает для удовлетворения потребностей в максимальных объемах, производитель заинтересован в экономии средств на организацию поставки. [8]

Обозначим  $F_{min}$  - минимальные суммарные расходы по доставке материальных средств для случая, когда спрос потребителей удовлетворен на минимальном уровне, равном  $b_j - s_j, j=1,2,\dots,m$ .

Степень уверенности в том, что предлагаемый вариант организации поставок экономически эффективен, будем характеризовать значением функции принадлежности затрат, соответствующих нечеткому числу  $F'$  - «расходы почти минимальны», описывающему цель, состоящую в их минимизации. [9]

Обозначим  $F(x)$  значение затрат, связанных с организацией поставки  $x$ :

$$F(x) = \sum_{i=1}^n \sum_{j=1}^m f_{ij} x_{ij},$$

Тогда, чем больше разность между затратами  $F(x)$  и величиной  $F_{min}$ , тем меньше степень уверенности в том, что план  $x$  экономически эффективен. [6]

Будем исходить из того, что  $F'$  - нечеткая величина, которой соответствует нечеткое треугольное число:

$$\langle -\infty, F_{min}, F_{max} \rangle,$$

где  $F_{max}$  — минимальные расходы, в случае, когда потребности клиентов удовлетворены в полном объеме  $-b_j, j = 1,2,\dots,m$ ; или (в случае недостатка запасов) таковы, что общая удовлетворенность потребителей максимальна. [11]

Функция принадлежности нечеткой цели, характеризующей эффективность поставки обозначим  $\mu_f(x)$ :

$$\mu_f(x) = \begin{cases} 0, & \text{нет допустимых решений,} \\ \frac{F(x) - F_{min}}{F_{max} - F_{min}}, & x - \text{допустимый план поставок} \end{cases} \quad (1)$$

Необходимо определить вариант поставки материальных средств, при котором степень уверенности (значение функции принадлежности) в том, что он эффективен по затратам и уровню обеспечения потребителей, максимальна.

### Математическая модель задачи логистики с нечеткими потребностями

Рассмотрим ограничения по удовлетворению спроса  $j$ -го потребителя:

$$\sum_{i=1}^n x_{ij} \geq b_j^*, j = 1, 2, \dots, m.$$

С учетом того, что  $b_j^*$  — нечеткие величины, эти ограничения задают нечеткие множества с функциями принадлежности  $\mu_j(x)$ :

$$\mu_j(x) = \begin{cases} 1, & \text{если } \sum_{i=1}^n x_{ij} \geq b_j, \\ \mu_{b_j^*} \left( \sum_{i=1}^n x_{ij} \right), & \text{если } \sum_{i=1}^n x_{ij} < b_j. \end{cases} \quad (2)$$

Оптимальный вариант поставки должен принадлежать всем нечетким множествам вида (2) и, кроме того, должен принадлежать нечеткому множеству цели с функцией принадлежности (1), следовательно, он должен принадлежать пересечению указанных нечетких множеств. Функция принадлежности этого пересечения имеет вид:

$$\mu_{\cap}(x) = \min\{\mu_f(x), \mu_1(x), \dots, \mu_m(x)\}.$$

Заметим, что функция  $\mu_{\cap}(x)$  характеризует степень уверенности в том, что план  $x$  эффективен по суммарным затратам и уровню обеспечения клиентов.[12] Необходимо максимизировать степень этой уверенности. С учетом сказанного получаем математическую модель задачи:

$$\left\{ \begin{array}{l} \min\{\mu_f(x), \mu_1(x), \dots, \mu_m(x)\} \rightarrow \max, \\ \sum_{i=1}^n x_{ij} \geq b_j^*, j = 1, 2, \dots, m, \\ \sum_{j=1}^m x_{ij} \leq a_i, i = 1, 2, \dots, n, \\ x_{ij} \geq 0, i = 1, 2, \dots, n, j = 1, 2, \dots, m. \end{array} \right. \quad (3)$$

### Анализ результатов работы

Для решения задачи (3) было разработано программное обеспечение на языке программирования С#, имеющее удобный графический интерфейс, с помощью которого проводился анализ результатов при различных исходных данных.

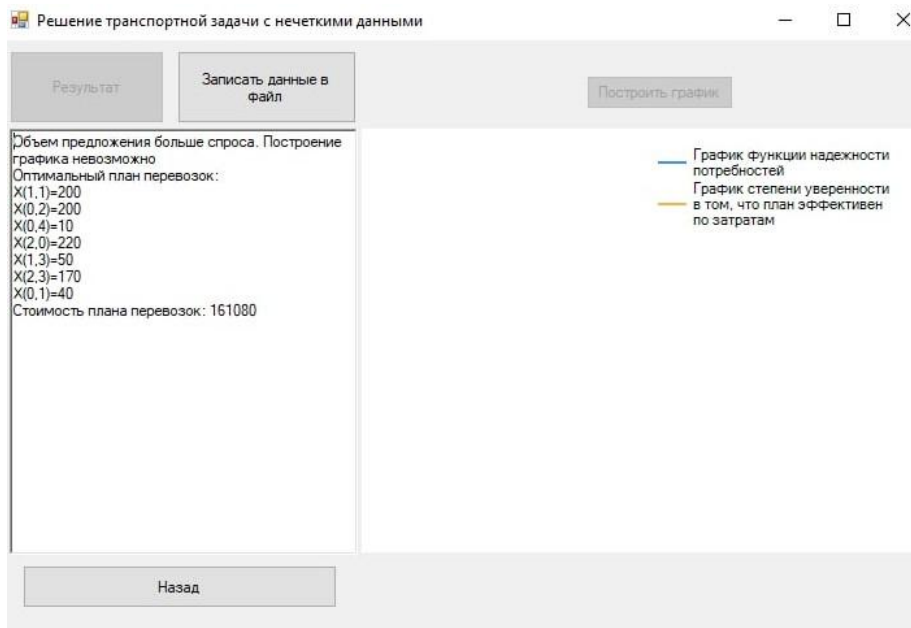
Процесс решения разбивается на шаги (число шагов можно задать в программе), каждый из которых заключается в решении транспортной задачи с определенным уровнем удовлетворения спроса потребителей, вычисляются значения надежности функции цели полученного решения. По этим значениям строится график, показывающий степень уверенности в том, что полученный план эффективен по затратам. Экспериментальным путем доказано, что чем меньше надежность функции цели полученного решения, тем менее выгодным является удовлетворение спроса потребителей.

Кроме того, программное обеспечение позволяет построить график функции удовлетворения потребностей, который отображает степень уверенности в том, что полученный план эффективен по уровню обеспечения потребителей. План тем лучше, чем больше потребностей возможно удовлетворить.

Точка пересечения этих графиков соответствует минимальному значению расходов при максимальном удовлетворении спроса потребителей, по ее координатам, определяется оптимальный план перевозок.

Произведем тестирование разработанного программного обеспечения на различных исходных данных и анализ полученных результатов.

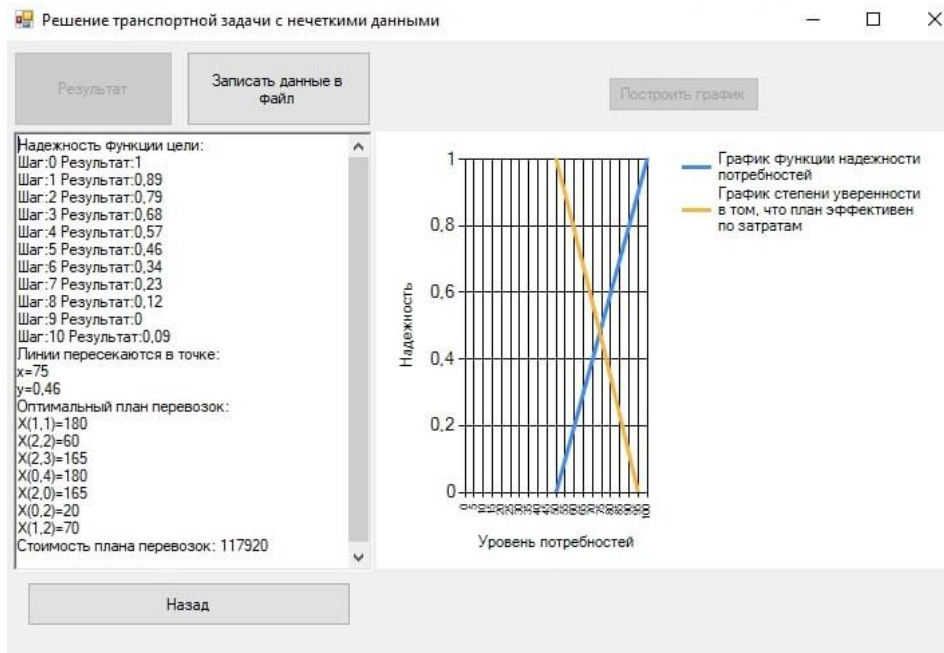
*Вариант 1.* Сумма произведенной продукции превышает 100% потребностей (рис.1).



**Рис.1. Результат работы при условии избыточного производства**

В случае избыточного производства (рис.1) нечеткая логика не применялась, происходило решение классической транспортной задачи. При изменении настроек ПО возможно решение этой задачи и с применением нечеткого подхода, который заключается не только в максимальном удовлетворении потребностей, но и минимизации издержек, в следствие чего не произойдет полного удовлетворения спроса потребителей, а будет найдено компромиссное решение, учитывающее заинтересованность в экономии ресурсов.

*Вариант 2.* Сумма произведенной продукции - 840 ед., количество продукции, заявленное на пунктах потребления, - 880 ед.



**Рис. 2. Результат тестирования ПО при условии нехватки произведенной продукции**

Анализ результатов (рис.2) показывает, что с увеличением уровня потребностей, надежность функции цели на каждом шаге уменьшается. Это объясняется тем, что чем больше потребностей мы хотим удовлетворить, тем дороже обойдется план поставок.

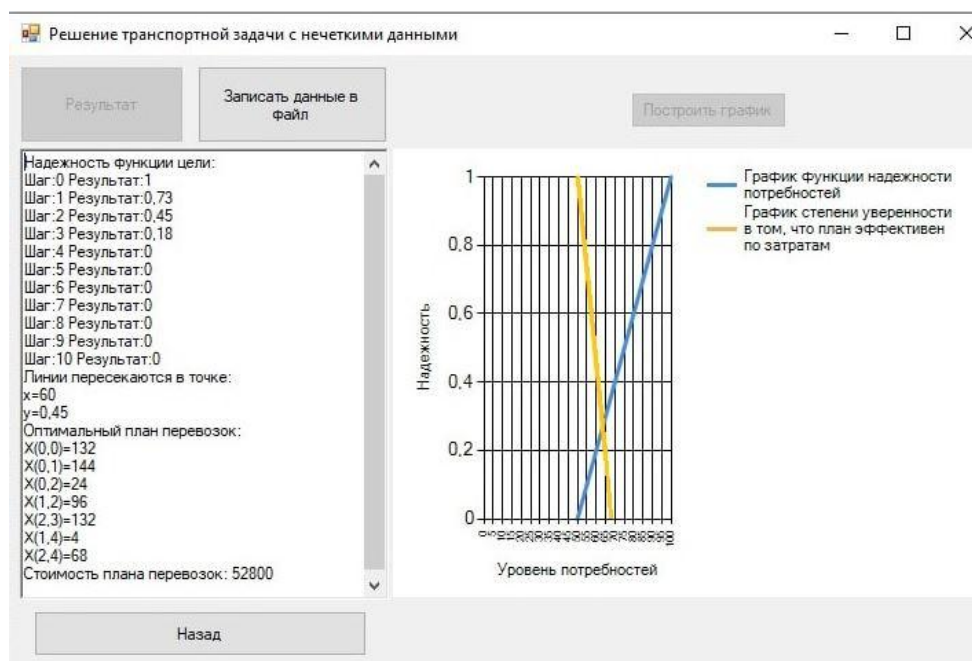
Координата точки пересечения графиков равна 75, что соответствует максимальному (в процентах) объему удовлетворения потребностей при минимальных затратах.

*Вариант 3.* Сумма произведенной продукции - 540 ед., количество продукции, заявленное на пунктах потребления, - 880 ед.

Так как предложение уменьшилось (рис.3), а спрос остался неизменным, то количество потребностей, которое мы сможем удовлетворить, тоже уменьшится по сравнению с предыдущим примером. Координата точки пересечения графиков равна 60. Что соответствует удовлетворению спроса на 60.

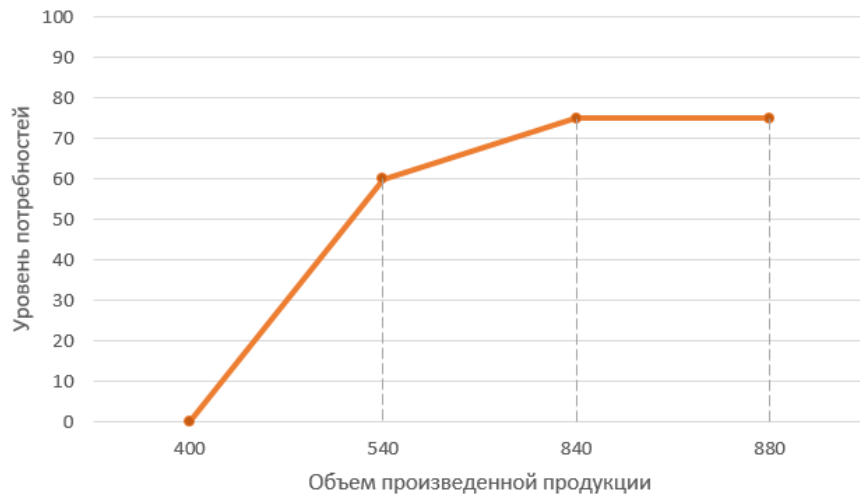
*Вариант 4.* Сумма произведенной продукции менее 50% от заявленного потребителем.

Поскольку в этом случае количество продукции, имеющееся на пунктах производства, не удовлетворяет даже минимальный уровень потребностей, на который согласны потребители, то данная задача не имеет решения. Программа выдала соответствующее сообщение.



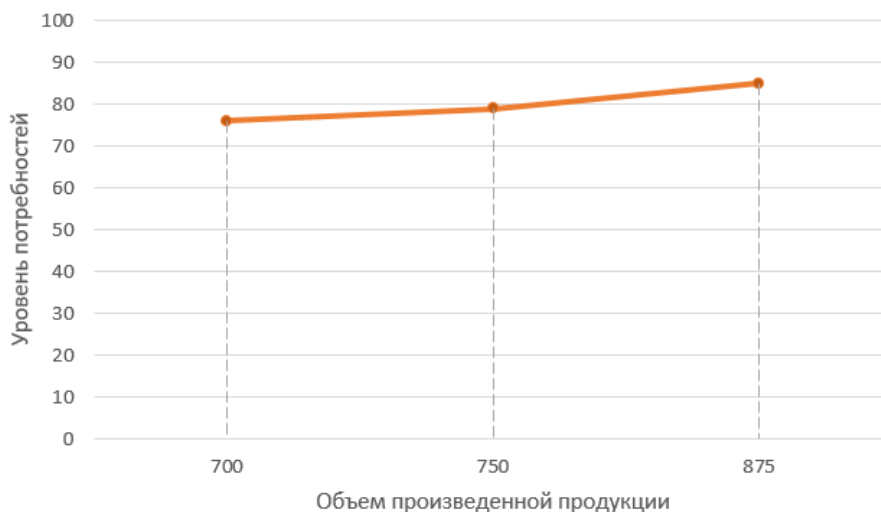
**Рис. 3. Результат тестирования ПО при увеличении допоставок продукции**

Для анализа полученных результатов построим график зависимости уровня удовлетворения потребностей в процентах от объема производства (рис.4). Объем спроса будем считать неизменным и равным 880 ед. Из графика видно, что с увеличением объема произведенной продукции, увеличивается и уровень удовлетворения потребностей. Максимальный уровень удовлетворения потребностей, который мы можем обеспечить, даже при приближении объема произведенной продукции к суммарному объему спроса (880ед.), будет равен 75 процентам, что объясняется стремлением минимизировать расходы на организацию плана поставок.



**Рис.4. График зависимости уровня потребностей от суммарного объема производства**

Увеличим минимальный уровень обеспечения, на который согласен каждый потребитель, с 50 до 70 процентов. Объем продукции, заявленный на пунктах потребления, так же оставим фиксированным – 880 ед. Построим график зависимости уровня удовлетворения потребностей в процентах от суммарного объема производства (рис.5).



**Рис. 5. График зависимости уровня потребностей от суммарного объема производства при увеличении минимального уровня обеспечения**

Наблюдаем, что уровень потребностей, который возможно удовлетворить, увеличился, по сравнению с предыдущим примером, когда минимальный уровень обеспечения составлял 50 процентов.

Таким образом, можно заметить, что уменьшение величины недопоставки продукции в пункты потребления влечет за собой увеличение степени уверенности в том, что потребитель удовлетворен.

### Вывод

В ходе работы было рассмотрено решение транспортной задачи с применением нечеткой логики, в которой потребности являются нечеткими величинами. По сравнению с обычным подходом, в котором исходные данные заданы определенно, рассмотренный нами подход более эффективен к решению реальных транспортных задач, поскольку он учитывает возможность недопоставок в пункты потребления, при этом оставляя потребителя удовлетворенным.

В дальнейшем планируется исследовать применение эвристических алгоритмов в решении транспортной задачи и проследить влияние эвристик на задачи с нечеткими данными.

#### Список литературы

1. **Косенко, О.В.** Разработка методов и алгоритмов решения многоиндексных распределительных задач в условиях неопределенности: дис. канд. тех. Наук. – Таганрог, 2017.- С.5-6.
2. **Ващекин, А.Н.** Применение математических методов теории нечетких множеств при моделировании принятия решений в экономической и правовой сфере // Статистика и Экономика.-2013.-№6.-С.18-21.
3. **Ghiaus, C.** Fuzzy Model and Control of a Fan-coil // Journal of Energy and Buildings. 2001. Vol. 33. P. 545–551.
4. **Круглов, В.В., Дли, М.И., Голунов, Р.Ю.** Нечеткая логика и искусственные нейронные сети. М.: Физматлит, 2001. 221 с.
5. Практическое применение нечеткой логики в задачах оптимизации [Электронный ресурс].-Режим доступа [https://studme.org/204548/matematika\\_himiya\\_fizik/prakticheskoe\\_primenenie\\_nechetkoy\\_logiki](https://studme.org/204548/matematika_himiya_fizik/prakticheskoe_primenenie_nechetkoy_logiki)
6. **Флегонтов, А.В.** Моделирование задач принятия решений при нечетких исходных данных: монография/ А.В. Флегонтов, В.Б. Вилков, А.К. Черных. – Санкт-Петербург: Лань, 2020. – С.54-64.
7. **Yager, R. R., Zadeh, L. A.** An Introduction to Fuzzy Logic Applications in Intelligent Systems. Boston: Kluwer Academic Publishers, 1992. 356 p.
8. **Zadeh, H. S.** Constrained multi-objective optimization of a fuzzy logic controller-applications in a non-linear system and a goal-seeking rover // Proc. of IEEE Conf. on Control Applications. 2003. V. 1. Istanbul, Turkey, 25 June 2003. P. 240–244.
9. **Шеври, Ф., Гели, Ф.** Нечеткая логика // Техническая коллекция. – 2009.- №31. – С. 4-6.
10. **Bellman, R. E., Zadeh, L. A.** Decision-making in fuzzy environment // Manag. Sci. 1970. 17, N 4. P. 141–160.
11. **Конюхов, А.Н., Дюбуа, А.Б., Сафошкин, А.С.** Основы теории нечетких множеств. Часть 2: учеб. пособие.- Рязань.- Рязан. гос. радиотехн. ун-т, 2018.- С. 7-10.



**РАЗРАБОТКА ПРИЛОЖЕНИЯ ДЛЯ ЗАДАЧ ОБРАБОТКИ  
ИЗОБРАЖЕНИЙ КЛЕТОК КРОВИ****П.А. Шагалова**

ORCID: 0000-0002-6676-4228 e-mail: polli-shagalova@yandex.ru

**Э.С. Соколова**

ORCID: 0000-0003-0860-2463 e-mail: essokolowa@gmail.com

**С.Н. Рындов**

ORCID: 0000-0002-8541-3757 e-mail: s.rindow@yandex.ru

**А.В. Бухнин**

ORCID: 0000-0003-3384-5248 e-mail: bukhnin@yandex.ru

Нижегородский государственный технический университет им. Р.Е. Алексева

В докладе представлены результаты разработки приложения, используемого для анализа изображений клеток крови, определения их характеристик - деформируемости, размеров, степени агрегации, с использованием технологий компьютерного зрения. Приложение включает несколько модулей, реализующие алгоритмы обработки изображений микроскопии крови. Для тестирования приложения использована реальная база данных изображений микроскопии крови, для увеличения размеров которых применены методы аугментации. Разработанное приложение может использоваться для обработки изображений микроскопии в области биологии, материаловедения, микроэлектроники. Микросервисная архитектура приложения позволяет встраивать новые решения, локально делать обновления, выбирать эффективные инструменты для обработки разного вида изображений. Существующие зарубежные решения очень дороги, что делает актуальным разработку алгоритмов анализа изображений микроскопии.

**Ключевые слова:** компьютерное зрение, обработка изображений, медицинское приложение, QT, OpenCV

**CREATING AN APPLICATION FOR TASKS OF BLOOD CELLS IMAGES PROCESSING****P.A. Shagalova, E.S. Sokolova, S.N. Ryndov**

Nizhny Novgorod state technical University n.a. R.E. Alekseev

**Purpose:** to describe the architecture of software development for practical and scientific use in the field of medicine.

**Design / Methodology / Approach:** an analysis was performed and the choice of a platform for creating software for medical image analysis was justified - the implementation of the software in C++ with QT Creator. Computer vision algorithms of the OpenCV library are used for image processing. To build the OpenCV library on the QT platform, the cross-platform utility CMake is used. The application includes the implementation of algorithms for determining the properties of blood cells, the analysis of aggregation processes, the determination of particle sizes on images of blood cells with the calculation of the Ferré diameter. Magnification methods were investigated to increase the sample size. For this purpose, the mdbloice/Augmentor library implemented in Python was used, which allows working with image masks. By applying magnification methods to the original image and the resulting mask, the efficiency of working on the magnification of the dataset was significantly increased without having to relabel the resulting images.

**Findings:** The article presents the results of automated processing of medical images that provide new information for diagnosis and scientific medical research.

**Research limitations/implications:** The research is limited by the medical image database. The architecture of the developed blood cell analysis application is scalable and allows the addition of new modules for specific research purposes.

**Originality/Value:** Currently, medical image analysis is performed using costly third-party solutions. Automation of microscopic image analysis algorithms based on artificial intelligence will improve the efficiency of object recognition in images. The developed application can be used for microscopy image processing in the fields of biology, material science, and microelectronics. The microservice architecture of the application allows you to include new software modules and provide users.

**Keywords:** computer vision, image processing, medical application, QT, OpenCV

## Актуальность и цель работы

Рынок программного обеспечения, предназначенного для анализа медицинских снимков, испытывает значительный рост в связи с интенсивным развитием технологий в области систем медицинской визуализации и анализа медицинских изображений. Разработка прикладных систем компьютерного зрения в области медицины, проведение исследований в области разработки эффективных алгоритмов обработки изображений, полученных при микроскопии крови, является актуальным направлением автоматизации биомедицинских исследований. [1] В области обработки изображений микроскопии крови преобладает ручной визуальный анализ, а ошибки в области лабораторной медицины, прямо влияющие на здоровье человека, стоят очень дорого. Автоматизация процесса, связанного с распознаванием сложных структурных образов на биомедицинских изображениях, определения их количественных характеристик значительно повысит эффективность и точность постановки диагноза, а также снизит затраты на лабораторные исследования. Разработанное и внедренное специализированное программное обеспечение в области решения трудоёмких медицинских задач позволит сотрудникам медицины не испытывать проблем с обработкой огромного количества медицинских данных, а больше сконцентрироваться на исследованиях и принятии важных решений, основываясь на результатах вычислений программы.

В настоящее время собраны большие базы данных (БД) изображений, полученных при микроскопии крови. В данном исследовании использовались реальные БД изображений недеформированных и деформированных эритроцитов и изображения тромбоцитов. Деформируемость эритроцитов в значительной степени определяет характер микроциркуляции крови, а потому имеет прямое отношение к диагностике и лечению многих заболеваний, в том числе таких распространенных, как нарушения мозгового кровообращения, сердечно-сосудистые заболевания, сахарный диабет, заболевания печени, желудочно-кишечного тракта и др. Лабораторные исследования и оценка агрегационной способности клеток крови очень актуальны в связи с потенциальной опасностью для жизни человека. Механизмы и последствия процессов агрегации эритроцитов и тромбоцитов значительно различаются с точки зрения влияния их на жизненно важные процессы организма [5]. Интенсивно проводятся медицинские исследования функциональных последствий динамичной агрегации-деагрегации эритроцитов как при патологических состояниях, так и в норме, когда наблюдается гиперагрегация и повышение вязкости крови, по-разному влияющая на транспорт кислорода в связанной с эритроцитарным гемоглобином форме. Определение степени агрегации тромбоцитов используется во всех анализах микрореологических свойств крови. При этом измерения времени агрегации тромбоцитов проводятся вручную, с помощью секундомера, фиксируя время, когда на стекле появляются отчетливые агрегаты [5,6], и далее визуально, с помощью микроскопа, определяют процентное распределение форм тромбоцитов.

Очевидно, что точность визуальных оценок структуры изображений, определение размеров, расположения характерных элементов в значительной степени зависит от квалификации лаборанта-диагноста и объема обрабатываемых данных. Процедура ручного анализа сложных по структуре изображений микроскопии представляет трудоемкий, однообразный процесс, что приводит к неточностям и ошибкам.

Таким образом, исследования в области анализа и автоматизации процессов распознавания изображений, полученных в результате микроскопии крови, являются актуальной задачей, тем более, что накоплены большие базы данных изображений клеток крови. Также интерес представляет задача динамического исследования процессов агрегации, т.е. обработки и анализа серии изображений, полученных в результате лабораторных исследований препаратов крови или физического моделирования процессов агрегации, что позволит лучше понять эти механизмы и их влияние на реологические свойства крови.

## Выбор инструментов для создания ПО

Для создания программного обеспечения, которое будет использоваться в практических и исследовательских целях в области анализа изображений микроскопии, важным аспектом является выбор среды разработки, обладающей необходимыми возможностями для реализации требуемых функции. Возможности среды разработки должны обеспечить создание эффективного и удобного для пользователя(медицинского работника) приложения. Среда разработки для построения комплексного медицинского приложения должна обладать кроссплатформенностью, чтобы производительность и удобство приложения не зависело от конкретной операционной системы. Должна быть совместимость элементов интерфейса с библиотекой компьютерного зрения OpenCV, а также поддержка различных компиляторов. Библиотека OpenCV содержит оптимизированные и широко используемые при обработки медицинских изображений алгоритмы, имеет открытый исходный код, свободно распространяется в академических и коммерческих целях, что позволяет применять её независимым разработчикам для создания собственных алгоритмов обработки изображений. [2]

Были рассмотрены популярные IDE, удовлетворяющие вышеперечисленным требованиям –Visual Studio и QT Creator.IDE Visual Studio имеет возможность разработки приложений под конкретную операционную систему, но всё же в большей степени предназначена для разработки приложений под Windows, работающих на.NET платформе. Среда разработки QTCreator,в свою очередь, делает акцент на возможности кроссплатформенного создания в этой среде приложений.[3] Ещё одним достоинством QT Creator по сравнению с Visual Studio является более удобная возможность выбора компилятора.

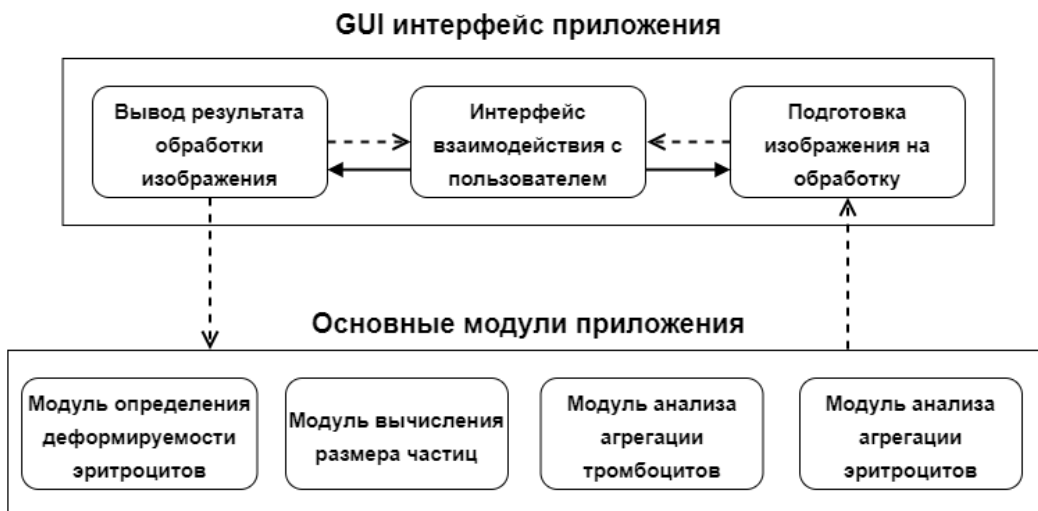
В результате проведенного анализа для разработки медицинского приложения выбрана свободная среда разработки QT Creator, позволяющая разрабатывать кроссплатформенные приложения для настольных, в том числе одноплатных компьютеров, и мобильных устройств. В качестве языка программирования, на котором написано приложение для обработки биомедицинских изображений, был выбран C++. Для обработки изображений выбрана библиотека алгоритмов компьютерного зрения OpenCV. Разработанное ПО на данном этапе представляет собой автономное решение, обладающее гибкостью и реализацией простых инструментов для обработки данных и оценки результатов. В то же время, при выборе платформы для разработки прикладного ПО, учитывалась возможность его интегрирования с медицинским оборудованием.

Для создания медицинского приложения в среде разработки QT Creator для полноценной разработки и исследования алгоритмов компьютерного зрения необходимо подключить библиотеку OpenCV, используя реализованную разработчиками поддержку сборки данной библиотеки под конкретную платформу. Для этого используется кроссплатформенная утилита автоматизации сборки CMake, позволяющая собрать необходимые файлы библиотеки под определённую среду разработки. При этом важно учесть, что в процессе сборки используется компилятор, который обязан поддерживать многопоточность программ. [4] В противном случае и без того сложный и довольно затратный по времени процесс сборки библиотеки OpenCV усложняется с возникновением ошибок. Это вызвано как раз оптимизацией алгоритмов библиотеки OpenCV, которые используют механизмы параллельного вычисления. После процесса сборки технологию OpenCV можно использовать в разработке на платформе QTCreator.

### Описание функциональности и архитектуры разработанного приложения

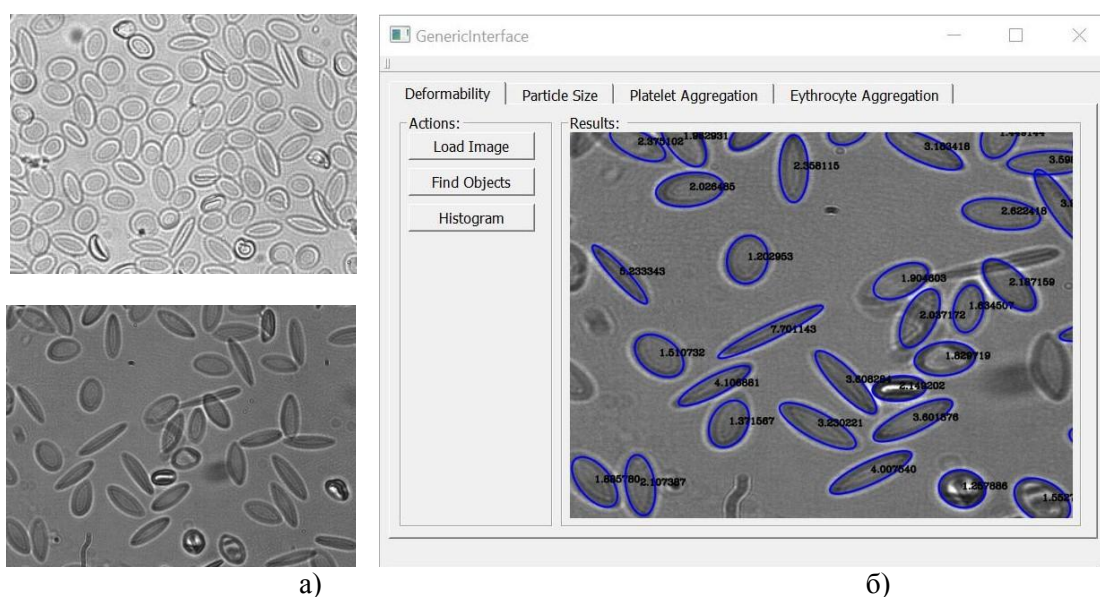
Разработанное приложение предоставляет возможность добавлять в систему новые функции, не нарушая структуры кода и не оказывая при этом влияния на существующую функциональность. Приложение включает следующие модули: модуль определения деформируемости эритроцитов; модуль определения размер частиц на изображении; модуль анализа агрегации эритроцитов и модуль агрегации тромбоцитов. Каждый модуль

представляет из себя отдельный алгоритм, который можно использовать для решения определённой задачи анализа изображений микроскопии крови. Архитектура приложения представлена на (рис.1).

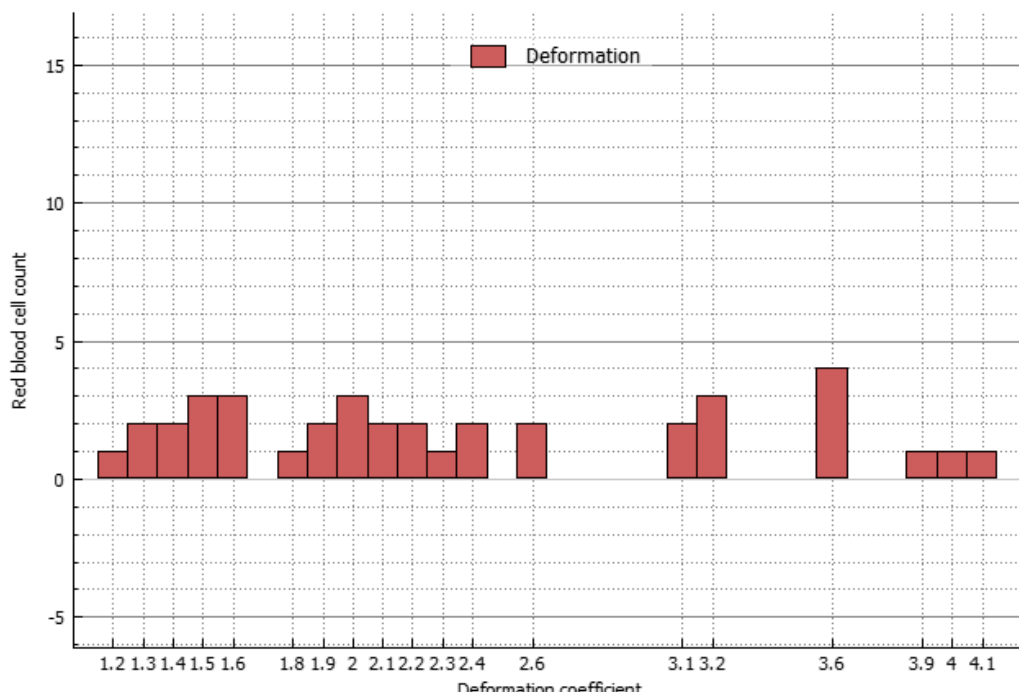


**Рис. 1. Архитектура разработанной системы**

На вход системы подаются изображения микроскопии крови, содержащиеся в соответствующей БД – снимки эритроцитов и тромбоцитов. Пользователь выбирает решаемую задачу – определение количественных характеристик эритроцитов, определение размера частиц, определение степени агрегации тромбоцитов и эритроцитов. В результате обработки изображения происходит распознавание и классификация объектов, а также вычисляются количественные характеристики клеток крови, значения которых выводятся в отдельный файл. Результат вычислений также можно представить в виде графика, показывающего распределение клеток крови по выбранному параметру. На (рис.2) представлены исходные изображения эритроцитов и результат их обработки, где выделены найденные эритроциты и выведены значения коэффициента их деформируемости. На (рис.3) представлен график, описывающий распределение клеток крови по вычисленному параметру, в данном случае – коэффициенту деформируемости эритроцитов.

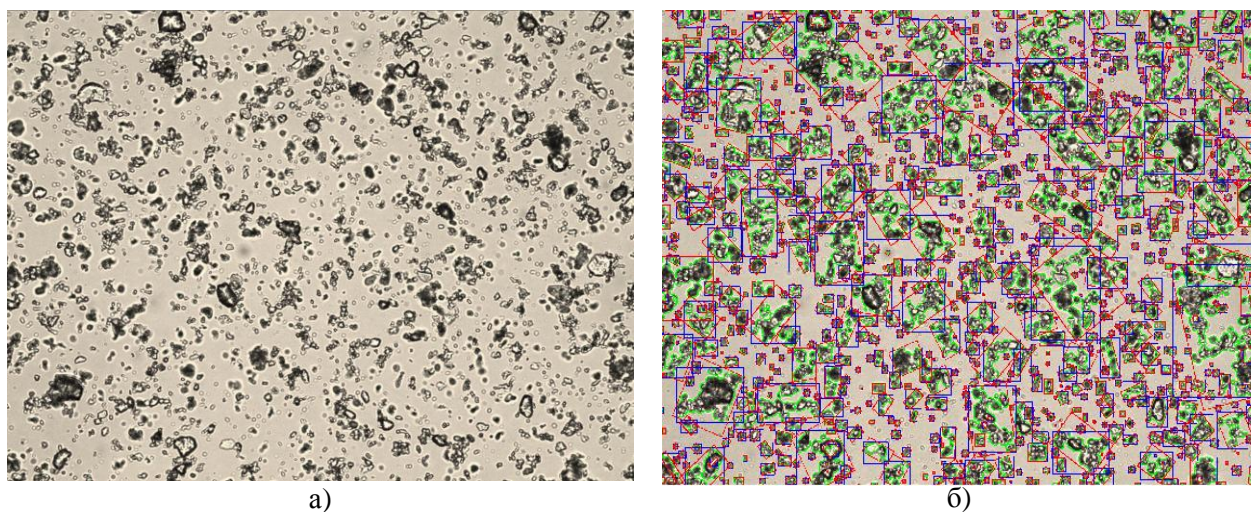


**Рис. 2. Анализ изображений эритроцитов а) исходные изображения эритроцитов, б) результат обработки изображения эритроцитов**



**Рис. 3. График распределения эритроцитов по степени их деформируемости**

Модуль вычисления размера частиц позволяет обработать визуально сложное изображение с большим количеством объектов, когда практически невозможно провести ручной анализ изображения для определения количества, концентрации частиц, их размеров, а также оценки степени их агрегации. На (рис.4) приведено изображение частиц препарата, подаваемое на вход системы, и результат обработки изображения, где выделены найденные частицы препарата. Для поиска частиц на изображении к исходному изображению применялся алгоритм бинаризации треугольника. К полученному бинаризованному изображению применяется морфологический фильтр и фильтр срединного размытия. Морфологический фильтр реализован через алгоритмы расширения и эрозии изображения. Для поиска частиц на изображении выполнены настройки алгоритмов фильтра срединного размытия, треугольной бинаризации, расширения и эрозии, которые входят в состав библиотеки OpenCV.



**Рис. 4. Результаты обработки изображения с целью вычисления размеров частиц препарата а) исходное изображение частиц препарата, б) результат обработки изображения с определением размеров частиц препарата**

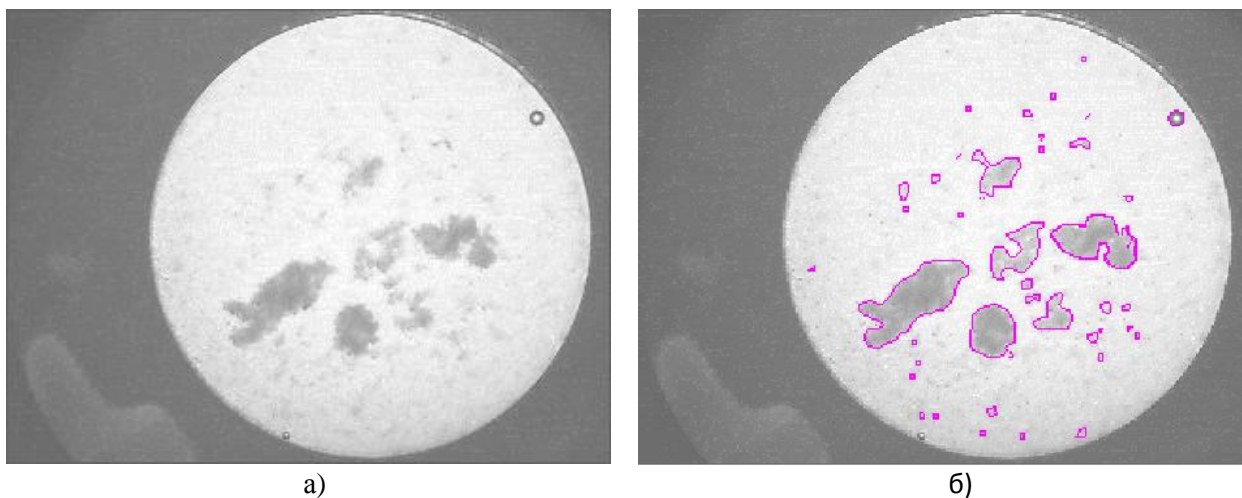


Затем к изображению применяется настроенный алгоритм поиска контуров, и на основе полученных данных для каждой частицы обрисовывается её контур, выделяя расположение частицы на изображении.

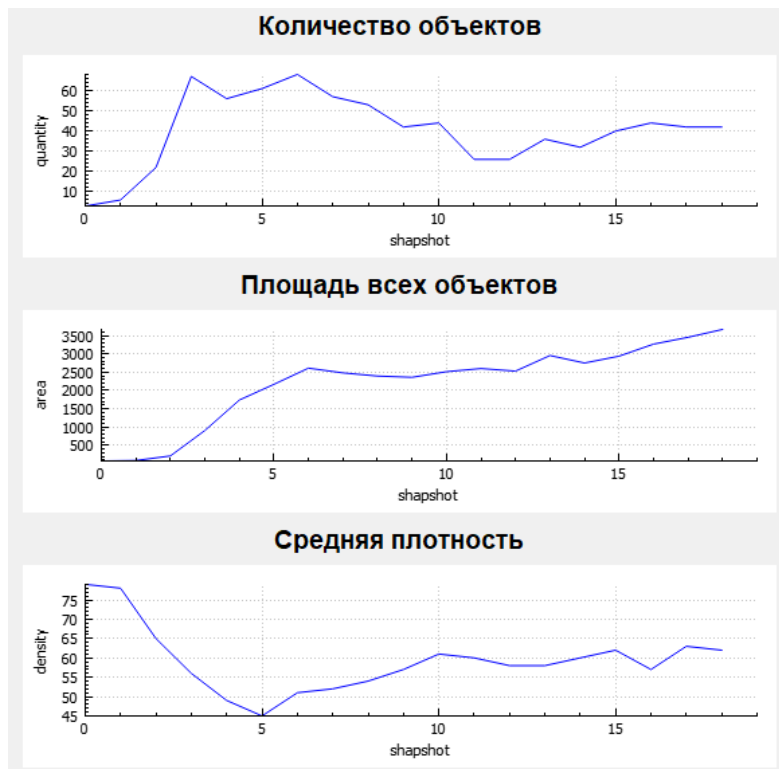
Функционал данного модуля не ограничивается обнаружением частиц препарата, он позволяет вычислить диаметр Ферре, который так же, как и контур, обрисовывается для каждой найденной частицы. Полученное изображение в формате Mat библиотеки OpenCV конвертируется в формат QImage платформы Qtcreator, после чего выводится в GUI интерфейс.

Модуль анализа процесса агрегации тромбоцитов позволяет определить на изображении агрегаты, подсчитать их количество, общую и среднюю площадь найденных объектов, а также плотность агрегата (рис.5). Для выделения агрегатов к исходному изображению применялся алгоритм бинаризации Брэдли, затем к исходному изображению применялся алгоритм пороговой бинаризации, и выполнялась дизъюнкция бинаризованных изображений. Описанные алгоритмы бинаризации используют алгоритм перевода цветных изображений в чёрно белое из библиотеки OpenCV. После применения алгоритма бинаризации, производится поиск контуров агрегатов с использованием алгоритма поиска контуров. На основе полученных данных вычисляется площадь и количество выделенных агрегатов. Полученное Mat изображение конвертируется в формат QImage, после чего так же выводится в GUI интерфейс.

Данные о количественных значениях параметров агрегации можно получить как в упрощённом численном эквиваленте, так и в расширенном, при помощи вывода соответствующих графиков, дающих больше диагностической информации о реологических свойствах крови (рис.6).

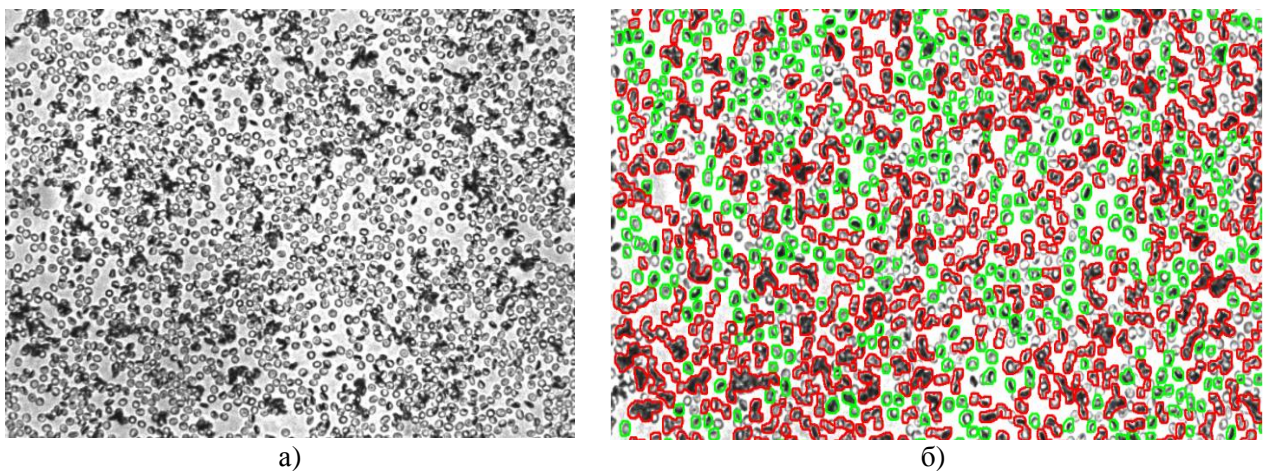


**Рис. 5. Результаты обработки изображений для определения оценки агрегации тромбоцитов) исходное изображение, б) Результат обработки изображения тромбоцитов с обозначением контуров**



**Рис. 6. Графики вычисленных данных агрегации тромбоцитов**

Модуль анализа процесса агрегации эритроцитов позволяет проанализировать сложное изображение и определить одиночные эритроциты и агрегаты (рис.7), выделяя на их на изображении соответствующим цветом (красные – агрегаты, зелёные – одиночные эритроциты). К исходному изображению применяются фильтр размытия по Гауссу и алгоритм бинаризации. Для получения информации о контурах эритроцитов изображение обрабатывается с помощью алгоритма водораздела, который основан на применении операции преобразования расстояния и нормализации, входящие в состав библиотеки OpenCV. После применения алгоритма водоразделак полученному изображению, для уменьшения шума применены алгоритмы расширения и эрозии. Затем применяется алгоритм поиска контуров, которые обрисовывают на изображении отдельные эритроциты и их агрегаты, затем результат выводится в GUI интерфейс.



**Рис.7. Результаты обработки изображений для нахождения эритроцитов и агрегатов а) Исходное изображение агрегатов и эритроцитов, б) Результат обработки изображения агрегатов и эритроцитов с обозначением контуров**

## Перспективы развития приложения

В настоящее время ведётся разработка альтернативного web интерфейса для медицинского приложения. Webинтерфейс позволяет упростить взаимодействие пользователя с приложением, когда взаимодействие будет происходить не через графический интерфейс, а через страницу в браузере. Так же, в разработке алгоритмов, больше не нужно будет учитывать особенности операционной системы пользователя.

В дальнейшем планируется проведение исследований в области изучения возможностей использования нейронных сетей для решения задач обработки изображений микроскопии крови, и интеграции в медицинское приложение алгоритмов на основе нейронных сетей.

Для этого планируется использовать технологии, работающие совместно с web интерфейсом. Такой технологией является Docker, которая позволяет «оборачивать» написанные на различных языках программирования решения в образ, который в дальнейшем будет запускаться и работать в изолированной среде, не влияющей на основную операционную систему. Такой способ упаковки и запуска позволяет подключить к медицинскому проекту модули, написанные на языке Python, что в свою очередь позволяет использовать алгоритмы нейронных сетей. Отдельные модули будут работать в отдельных контейнерах и выводить результат обработки медицинских изображений в web интерфейс. Результаты обработки изображений так же могут загружаться на сервер, доступ к которому будут иметь только авторизованные пользователи.

Web интерфейс также позволяет добавить к приложению базу данных и систему авторизации пользователей, обеспечивая индивидуальный подход к пользователю. Результаты обработки изображений, информация для сбора медицинской статистики могут храниться в базе данных, доступ к которой получит только авторизованный пользователь. База данных предоставляет возможность решения проблемы использования нейронных сетей, т.е. создания и использования большого датасета для обучения сетей. Актуальные медицинские изображения могут добавляться пользователями разработанного медицинского приложения. Обучение нейронных сетей на актуальных медицинских изображениях позволит улучшить качество обработки изображений.

По итогу разработка альтернативного web интерфейса открывает новые возможности как для разработчиков, так и для пользователей приложения. Гибкая web система позволяет создавать новые программные решения конкретной медицинской задачи, без привязки к определённому языку программирования, при этом риск потери производительности работы приложения в следствии недостаточно мощной системы пользователя сводится к минимуму, так как основной вычислительный процесс будет происходить на сервере, а не на компьютере пользователя.

## Библиографический список

1. ИТ в медицине: от информатизации к цифровой трансформации [Электронный ресурс].- Режим доступа <https://www.iksmmedia.ru/articles/5863658-IT-v-medicine-ot-informatizacii-k.html>
2. OpenCV Wiki [Электронный ресурс].-Режим доступа <https://github.com/opencv/opencv/wiki>
3. About QT [Электронный ресурс].-Режим доступа [https://wiki.qt.io/About\\_Qt](https://wiki.qt.io/About_Qt)
4. OpenCV withQT [Электронный ресурс].-Режим доступа [https://wiki.qt.io/OpenCV\\_with\\_Qt](https://wiki.qt.io/OpenCV_with_Qt)
5. Обратимая агрегация эритроцитов при течении крови с возрастающей и с уменьшающейся скоростью / Соколова И.А., Качалова А.В., Шахназаров А.А., Кошелёв В.Б. // Журнал «Тромбоз, гемостаз и реология». – Москва, том 74, № 2, с. 54-62
6. **Медведев, И.Н., Завалишина, С.Ю., Краснова, Е.Г., Кутафина, Н.В.** Методические подходы к оценке агрегации и поверхностных свойств тромбоцитов и эритроцитов // Фундаментальные исследования. – 2014. – № 10-1. – С. 117-120



## СРАВНИТЕЛЬНАЯ РЕАЛИЗАЦИЯ АЛГОРИТМОВ ПОВЫШЕНИЯ РАЗРЕШЕНИЯ ИЗОБРАЖЕНИЙ

Д. С. Бирюков,

ORCID0000-0003-4756-0294 e-mail: danbir@mail.ru

О. Н. Корелин,

ORCID 0000-0002-5928-6045 e-mail: o\_korelin@mail.ru

Д. А. Кобляков

ORCID 0000-0002-7980-5801 e-mail: d.koblyakov@gmail.com

Нижегородский государственный технический университет им. Р.Е. Алексева

В настоящее время большое значение имеет повышение качество фото и видео. Чем выше разрешение, тем качественнее картинка, тем резче и четче контуры. Существуют математические методы и алгоритмы позволяющие повысить разрешение, они относятся к классу методов Super-Resolution (SR). Наилучший эффект даёт использование нейронных сетей. В докладе предлагается реализация алгоритмов SR на разных платформах с использованием библиотек OpenCV и OpenVINO. Приводится сравнение времени выполнения методов SR на различных устройствах (CPU(x86), CPU(arm), GPU, Intel Neural Compute Stick 2). Имеется описание по настройке этих устройств, которое может быть использовано в методических целях. Особенно выделена работа с одноплатным компьютером RaspberryPi и специализированным нейросетевым сопроцессором IntelNeuralComputeStick.

**Ключевые слова:** Super-Resolution, OpenVINO, OpenCV, одноплатный компьютер Raspberry Pi, Intel Neural Compute Stick, нейронные сети, повышение разрешения изображений.

## COMPARATIVE IMPLEMENTATION OF IMAGE RESOLUTION UPGRADE ALGORITHMS

D.S. Birukov, O.N. Korelin, D.A. Koblyakov

Nizhniy Novgorod state technical university n.a. R.E. Alekseev

**Purpose:** The purpose of this article is to provide possible approaches to run inference of super-resolution neural networks. There are provide few frameworks and devices, described mean inference time and also gives guide to set up all this approaches.

**Design/methodology/approach:** The paper use inference metrics to compare different devices for performance in super-resolution task. There also provides description of setup, provides recommendations and requirements.

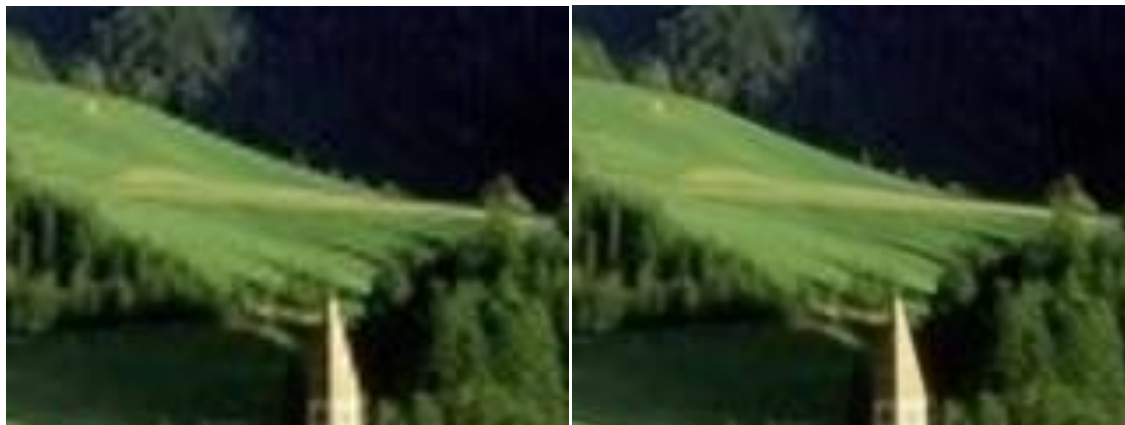
**Findings:** Compare of different devices in task of super-resolution. Provides description with requirements and recommendations to setup different approaches. Described common approach of usage frameworks with super-resolutions models with single-board computer.

**Research limitations/implications:** The present article provides a starting-point for further usage of super-resolutions models on singel-board computer.

**Originality/value:** Provided approach could be used on all systems that used super-resolution on single board computer. It's helps to select with devices for the system being created. It's also may be used in methodological goals.

**Keywords:** Super-Resolution, OpenVINO, OpenCV, Raspberry Pi Single Board Computer, Intel Neural Compute Stick, neural network, Image Upscaling

Наиболее простыми в реализации являются методы ближайших соседей (который использует для увеличения значение ближайшего пикселя из исходного изображения). Однако даже при кратном увеличении исходного изображения данные методы вносят весомые изменения, видимые не вооружённым взглядом, а потому спектр применения данных методов крайне ограничен. В последнее же время все большим успехом пользуются методы, основанные на использовании свёрточных нейронных сетей.



**Рис.1. Сравнение бикубической интерполяции (левое изображение) и Super Resolution (правое изображение)**

Эти алгоритмы используются для борьбы с шумом:



**Рис.2. Пример борьбы с шумом (левое изображение с шумом)**

Существуют как однокадровые, так и многокадровые варианты SR. Многокадровый SR использует так называемые субпиксельные сдвиги между несколькими изображениями с низким разрешением одной и той же сцены. Он создает изображение с улучшенным разрешением, объединяющее информацию со всех изображений с низким разрешением, а созданные изображения с более высоким разрешением лучше описывают сцену. Однокадровые методы SR пытаются увеличить изображение без размытия. Эти методы используют другие части изображений с низким разрешением или другие несвязанные изображения, чтобы угадать, как должно выглядеть изображение с высоким разрешением. Алгоритмы, так же, могут быть разделены по принципу исследования, т.е. проводимые в частотной или пространственной областях. Первоначально методы Super-resolution хорошо работали только на изображениях в оттенках серого, но исследователи нашли способы адаптировать их к цветным изображениям камеры. Недавно также было показано использование сверхразрешения для 3D-данных.

Во многом вопрос использования той или иной нейронной сети упирается во время выполнения на конечном устройстве. В данной работе рассматриваются способы уменьшения

времени выполнения нейронных сетей для повышения разрешения изображений на различных платформах.

Наиболее частыми аппаратными платформами для использования нейронных сетей являются CPU(central processing unit), GPU(graphics processing unit) и различные встраиваемые системы. Для проведения экспериментов имелись следующие платформы:

- CPU - Intel Core i5-2400 / Intel Core i5 - 10600k
- GPU - GEFORCE GTX 1060 / UHD 630
- Одноплатный компьютер Raspberry Pi 3 b+

В качестве программных платформ для работы с алгоритмами Super-resolution были выбраны OpenCV и TensorFlow.

OpenCV — это open source библиотека компьютерного зрения, которая предназначена для анализа, классификации и обработки изображений. Широко используется в таких языках как C, C++, Python и Java. В данной библиотеке имеется ряд встроенных моделей Super-resolution, что особенно актуально в данном исследовании. Так же OpenCV имеет поддержку ARM архитектуры, что особенно хорошо для запуска нейронных сетей на одноплатном компьютере Raspberry Pi. Если будет понятно, что текущих мощностей Raspberry Pi недостаточно можно будет перенести часть вычислений на Intel Neural Compute Stick.

TensorFlow — открытая программная библиотека для машинного обучения, разработанная компанией Google для решения задач построения и тренировки нейронной сети с целью автоматического нахождения и классификации образов, достигая качества человеческого восприятия. Применяется как для исследований, так и для разработки собственных продуктов Google. Основной API (Application Programming Interface) для работы с библиотекой реализован для Python. TensorFlow может работать на многих параллельных процессорах, как CPU, так и GPU, опираясь на архитектуру CUDA (Compute Unified Device Architecture) для поддержки вычислений общего назначения на графических процессорах.

Raspberry Pi 3 b+ это третье поколение одноплатных компьютеров малого размера, изначально разработанных как бюджетная система для обучения информатике, но позже получивших более широкое применение и известность. Разрабатывается британской компанией Raspberry Pi Foundation.

Данный одноплатный компьютер поддерживает работу множества операционных систем (Raspberry Pi OS(на основе Debian), Ubuntu, Windows 11). Рекомендованной для использования является Raspberry Pi OS.

Raspberry Pi не имеет видеокарты от компании Nvidia и не поддерживает CUDA, это накладывает ряд ограничений на использование фреймворков предназначенных для обучения моделей нейронных сетей. Основными фреймворками для работы с этим одноплатным компьютером являются OpenVINO и OpenCV. Причем OpenVINO хорошо подходит для работы с нейронными сетями, предоставляя высокую скорость исполнения. OpenCV в свою очередь скорее нацелен на выполнение общих задач компьютерного зрения.

В состав библиотеки OpenCV входит множество функций, которые могут понадобиться при решении задач компьютерного зрения, начиная от задач связанных с сохранением и чтением из памяти изображений заканчивая реализованными алгоритмами нахождения оптического потока изображений.

Так же в состав OpenCV входит модуль для повышения разрешения изображений - cv2.dnn\_superres. Однако данный модуль не присутствует в основной репозитории библиотеки, а находится в репозитории opencv\_contrib, который предназначен для нового функционала, не пригодного для включения в основной проект.

Установка OpenCV с модулем opencv\_contrib компьютер под управлением операционной системы на базе ядра Linux с процессором на архитектуре x86 не представляет особых сложностей. В открытом доступе имеется построенная версия библиотеки. Так например можно установить с помощью пакетного менеджера Python – pip:

```
python3 -m pip install opencv-contrib-python
```

Установка OpenCV с модулем `opencv_contrib` на компьютер под управлением операционной системы на базе ядра Linux с процессором на архитектуре ARM имеет ряд трудностей. В открытом доступе недоступна простроенная версия библиотеки. Таким образом необходимо скомпилировать данную библиотеку из исходного кода самостоятельно. В дальнейшем будет рассмотрена установка на одноплатный компьютер Raspberry Pi. Данная инструкция должна подходить и для других устройств на базе ARM процессоров. Отдельно надо обратить внимание на то, что для установки OpenCV на Raspberry рекомендуется иметь карту памяти, на которую установлена система объёмом не менее 16 гигабайт.

Перед сборкой OpenCV необходимо провести установку дополнительных пакетов:

- `sudo apt-get update && sudo apt-get upgrade`
- `sudo apt-get install build-essential cmake pkg-config`
- `sudo apt-get install libjpeg-dev libtiff5-dev libjasper-dev libpng-dev`
- `sudo apt-get install libavcodec-dev libavformat-dev libswscale-dev libv4l-dev`
- `sudo apt-get install libxvidcore-dev libx264-dev`
- `sudo apt-get install libfontconfig1-dev libcairo2-dev`
- `sudo apt-get install libgdk-pixbuf2.0-dev libpango1.0-dev`
- `sudo apt-get install libgtk2.0-dev libgtk-3-dev`

После установки дополнительных пакетов необходимо клонировать на устройства два репозитория (`opencv` и `opencv_contrib`):

- `git clone https://github.com/opencv/opencv.git`
- `git clone https://github.com/opencv/opencv_contrib.git`

Во время компиляции проекта используется много оперативной памяти и возможно, что встроенной памяти не хватит, поэтому рекомендуется увеличить размер `swap-file` (файл подкачки) до двух гигабайт. Для этого необходимо выполнить следующую команду:

- `sudo nano /etc/dphys-swapfile`
- Изменить размер до 2000

После изменений необходимо перезагрузить Raspberry. Когда перезагрузка закончена можно приступать к установке OpenCV. Для этого необходимо перейти в директорию с клонированным репозиторием:

- `cd opencv`
- Создание папки для компиляции `mkdir build`
- `cd build`

Далее необходимо запустить CMake для конфигурации компиляции фреймворка. Во время работы CMake может вернуть несколько сообщений об ошибке, на это можно не обращать внимания:

- `make -D CMAKE_BUILD_TYPE=RELEASE -D CMAKE_INSTALL_PREFIX=/usr/local -D OPENCV_EXTRA_MODULES_PATH=../opencv_contrib/modules -D ENABLE_NEON=ON -D ENABLE_VFPV3=ON -D BUILD_TESTS=OFF -D INSTALL_PYTHON_EXAMPLES=OFF -D OPENCV_ENABLE_NONFREE=ON -D CMAKE_SHARED_LINKER_FLAGS=-latomic -D BUILD_EXAMPLES=OFF .`

Запускаем построение библиотеки. Это может занять много времени, если будут какие-то критичные места, можно попробовать запустить `make` с флагом `-j1` и подождать некоторое время. В данном случае на то чтобы собрать библиотеку ушло около 12 часов:

- `make -j4`

Если предыдущий шаг прошел успешно, необходимо закончить установку:

- `sudo make install`

После завершения установки можно проверить успешность действий, запустив в интерпретаторе `python3` команду `import cv2`.

Оценка производительности нейронных сетей в библиотеке OpenCV производилась с использованием 4 нейронных сетей

#### EDSR

- Размер модели: ~38,5 МБ
- Доступны обученные варианты с коэффициентом увеличения: x2, x3, x4
- Преимущество: высокая точность
- Недостаток: медленная и имеет большой размер

#### ESPCN

- Размер модели: ~ 100 КБ
- Доступны обученные варианты с коэффициентом увеличения: x2, x3, x4
- Преимущество: высокая скорость работы и малое количество параметров при достаточном уровне точности

- Недостаток: показывает наихудшую точность из представленных моделей

#### FSRCNN

- Размер модели: ~ 40 КБ
- Доступны обученные варианты с коэффициентом увеличения: x2, x3, x4
- Преимущество: быстрая и легкая, при этом способная выдавать достаточно высокую точность

- Недостаток: проигрывает в точности некоторым своим конкурентам

#### LapSRN

- Размер модели: от 1 до 5 МБ
- Доступны обученные варианты с коэффициентом увеличения: x2, x4, x8
- Преимущество: Модель может выполнять многомасштабное увеличение разрешения с одним прямым проходом

- Недостаток: низкая производительность и точность

В ходе работы была проведена серия тестов и получены показатели времени работы нейронных сетей на тех или иных аппаратных платформах, результаты приведены в следующей таблице, все тесты проводились при разрешении входного изображения 100 на 100 пикселей и коэффициенте увеличения x4:

Таблица 1

Сравнение времени работы нейронных сетей на разных аппаратных платформах

	i5 2400+OpenCV (сек.)	gtx1060 + Tensorflow (сек.)	ARM Cortex- A53 (Raspberry Pi)+OpenCV (сек)	i5 10600k+OpenCV (сек.)
EDSR	93.8	1.01	2015.82	1.55
ESPCN	0.13	0.01	1.81	0.0052
FSRCNN	0.197	0.033	1.91	0.0077
LapSRN	6.43	0.89	109.6	0.14

Как можно увидеть - использование процессора Raspberry для исполнения нейронных сетей возможно, но время исполнения слишком велико и подходит не для всех систем. Так например для работы в видеопотоком с частотой кадров - 30 fps, требуется чтобы время выполнения было меньше 33 миллисекунд. Так же проблемой для использования OpenCV является ограниченность в использовании только 4 моделей.

Для решения вышеизложенных проблем можно использовать дополнительного вычислительного устройства Intel Neural Compute Stick 2 (NCS2) под управлением фреймворка



OpenVINO. Данное устройство специально создано для того чтобы повысить эффективность и быстродействие алгоритмов путём переноса на него вычисления нейронных сетей.



**Рис.3. Внешний вид и подключение Intel Neural Compute Stick 2**

Для работы с модулем необходимо установить фреймворк OpenVINO. Данный фреймворк разрабатывался специально для запуска нейронных сетей на исполнение с использованием устройств компании Intel, в том числе и Intel Neural Compute Stick 2.

OpenVINO - это бесплатный фреймворк, облегчающий оптимизацию моделей нейронных сетей и последующее развертывание на устройствах производства компании Intel. Он состоит из множества компонент. Начиная от оптимизатора моделей и среды для исполнения нейронных сетей, заканчивая открытой коллекцией нейронных сетей. OpenVINO поддерживает множество форматов для обучения нейронных сетей. Что позволяет конвертировать и оптимизировать обученные модели в формат использующийся в OpenVINO.

1. Для установки OpenVINO на Ubuntu (на ПК с процессором x86) необходимо установить OpenCL

- `sudo add-apt-repository ppa:intel-ocl/intel-ocl`
- `sudo apt update`
- `sudo apt install libze-loader libze-intel-gpu`
- `sudo apt install intel-ocl-icd`

2. По ссылке загрузить рекомендованный установщик

[https://www.intel.com/content/www/us/en/developer/tools/openvino-toolkit/download.html?operatingsystem=linux&distributions=webdownload&version=2021%204.2%20LTS%20\(latest\)&options=online](https://www.intel.com/content/www/us/en/developer/tools/openvino-toolkit/download.html?operatingsystem=linux&distributions=webdownload&version=2021%204.2%20LTS%20(latest)&options=online)

3. Распаковать `tar -xvzf l_openvino_toolkit_p_<version>.tgz`

4. Запустить установщик `sudo ./install_GUI.sh`

5. Следуя инструкции устанавливаем OpenVINO

6. Перейти в директорию с установленным OpenVINO `cd /opt/intel/openvino_2021/install_dependencies`

7. Установить зависимости `sudo -E ./install_openvino_dependencies.sh`

8. Исправить `.bashrc` файл ( `vi ~/.bashrc` ) добавив в конец файла строчку `source /opt/intel/openvino_2021/bin/setupvars.sh`

9. Установить зависимости для `model_optimizer`. Перейдя в директорию `cd /opt/intel/openvino_2021/deployment_tools/model_optimizer/install_prerequisites` и запустив установку зависимостей `sudo ./install_prerequisites.sh`

Для установки OpenVINO на Raspberry Pi:

1. Необходимо в OpenVINO Toolkit packages storage найти нужную версию пакета по такому шаблону `l_openvino_toolkit_runtime_raspbian_p_<version>.tgz` и загрузить её на Raspberry в папку Downloads и перейти в эту папку

- `cd Downloads`

- wget

[https://storage.openvinotoolkit.org/repositories/openvino/packages/2021.4.1/l\\_openvino\\_toolkit\\_runtime\\_raspbian\\_p\\_2021.4.689.tgz](https://storage.openvinotoolkit.org/repositories/openvino/packages/2021.4.1/l_openvino_toolkit_runtime_raspbian_p_2021.4.689.tgz) (команда будет отличаться для другой версии пакета)

2. Необходимо создать папку, в которую будет установлен OpenVINO  
`sudo mkdir -p /opt/intel/openvino_2021`
3. Теперь необходимо распаковать архива  
`sudo tar -xvf l_openvino_toolkit_runtime_raspbian_p_<version>.tgz --strip 1 -C /opt/intel/openvino_2021`
4. Для использования OpenVINO необходимо определить некоторые переменные окружения  
`source /opt/intel/openvino_2021/bin/setupvars.sh`
5. Необходимо обновить файл .bashrc  
`echo "source /opt/intel/openvino_2021/bin/setupvars.sh" >> ~/.bashrc`
6. В результате инициализации можно увидеть следующее сообщение от OpenVINO:  
`[setupvars.sh] OpenVINO environment initialized`
7. Дальнейшие шаги необходимо выполнять только если планируется использовать Intel Neural Compute Stick. Необходимо добавить текущего пользователя в группу users  
`sudo usermod -a -G users "$(whoami)"`
8. Настроить переменные окружения  
`source /opt/intel/openvino_2021/bin/setupvars.sh`
9. Установить правила USB для Neural Compute Stick.  
`sh /opt/intel/openvino_2021/install_dependencies/install_NCS_udev_rules.sh`
10. Вставить Intel Neural Compute Stick 2(NCS2) в USB порт Raspberry

На этом настройка OpenVINO на Raspberry закончена. Стоит предупредить, что OpenVINO состоит из множества компонент (Inference Engine, Model Optimizer, OpenCV и т.д). Данная установка использует заранее построенный проект лишенный многих компонент. Если имеется необходимость в этих компонентах, устанавливаемых на Raspberry - необходимо либо самостоятельно построить их на Raspberry, либо установить их на host-PC.

Для проверки работоспособности системы был использован предоставленные OpenVINO пример. В данной части будет использоваться компонента Open Model Zoo. Она предоставляет доступ ко множеству подготовленных нейронных сетей. Эту компоненту можно найти на GitHub [https://github.com/openvinotoolkit/open\\_model\\_zoo](https://github.com/openvinotoolkit/open_model_zoo)

Порядок запуска примера следующий:

1. Переходим в домашнюю директорию и создаём папку build
  - `cd ~`
  - `mkdir build && cd build`
2. Запускаем компиляцию примера
  - `cmake -DCMAKE_BUILD_TYPE=Release -DCMAKE_CXX_FLAGS="-march=armv7-a" /opt/intel/openvino_2021/deployment_tools/inference_engine/samples/cpp`
  - `make -j2 object_detection_sample_ssd`
3. Клонировать репозиторий Open Model Zoo, устанавливаем зависимости, и загружаем модель, используя утилиту от OpenVINO
  - `git clone --depth 1 https://github.com/openvinotoolkit/open_model_zoo`
  - `cd open_model_zoo/tools/downloader`
  - `python3 -m pip install -r requirements.in`
  - `python3 downloader.py --name face-detection-adas-0001`
4. Запустить пример на исполнение с использованием NCS2 (MYRIAD) в качестве изображения рекомендуется использовать любое фото с лицом человека. В результате появится файл out\_0.bmp с квадратом вокруг всех найденных лиц
  - `./armv7l/Release/object_detection_sample_ssd -m <path_to_model>/face-detection-adas-0001.xml -d MYRIAD -i <path_to_image>`
5. В результате работы должно появиться примерно следующее сообщение

6.

```
[142,1] element, prob = 0.0170898 (843,200)-(901,265) batch id : 0
[ INFO ] Image out_0.bmp created!
[ INFO ] Execution successful

[ INFO ] This sample is an API example, for any performance measurements please use the dedicated benchmark_app tool
pi@raspberrypi:~/build $
```

**Рис.4. Результаты проверки работы NCS2**

В Open Model Zoo имеется несколько примеров нейронных сетей повышающих разрешение. Например: single-image-super-resolution-1032. Для загрузки этой модели необходимо воспользоваться утилитой downloader из Open Model Zoo

- cd open\_model\_zoo/tools/downloader
- python3 downloader.py --name single-image-super-resolution-1032

Должно появиться следующее сообщение:

```
#####| Downloading single-image-super-resolution-1032 |#####
===== Downloading super_res_net/intel/single-image-super-resolution-1032/FP32/single-image-super-resolution-1032.xml
... 100%, 80 KB, 770 KB/s, 0 seconds passed
===== Downloading super_res_net/intel/single-image-super-resolution-1032/FP32/single-image-super-resolution-1032.bin
... 100%, 116 KB, 1267 KB/s, 0 seconds passed
===== Downloading super_res_net/intel/single-image-super-resolution-1032/FP16/single-image-super-resolution-1032.xml
... 100%, 80 KB, 2839 KB/s, 0 seconds passed
===== Downloading super_res_net/intel/single-image-super-resolution-1032/FP16/single-image-super-resolution-1032.bin
... 100%, 58 KB, 619 KB/s, 0 seconds passed
===== Downloading super_res_net/intel/single-image-super-resolution-1032/FP16-INT8/single-image-super-resolution-1032.xml
... 100%, 162 KB, 1111 KB/s, 0 seconds passed
===== Downloading super_res_net/intel/single-image-super-resolution-1032/FP16-INT8/single-image-super-resolution-1032.bin
... 100%, 30 KB, 478 KB/s, 0 seconds passed
```

**Рис.5. Результаты загрузки модели**

После того как модель была загружена ее можно использовать в своей программе. С использованием данной сети были сняты показатели производительности (Табл.2)

**Таблица 2**

**Показатели производительности**

	i5 10600k CPU (сек.)	UHD 630 iGPU (сек.)	NCS2 (сек.)
single-image-super-resolution-1032	0.1535	0.1016	2.3897

Из (Табл.2) видно, что использование встроенной графической карты в современных процессорах позволяет существенно улучшить время работы нейронных сетей. К сожалению на данный момент не получилось собрать данные о времени работы нейронной сети на процессоре одноплатного компьютера Raspberry Pi, судить о повышении времени работы в данном случае можно лишь теоретически соотнося таблицы (Табл. 1 и Табл. 2).

По окончании работы можно сформулировать следующие задачи:

- Запустить модель single-image-super-resolution-1032 с помощью OpenVINO. Для этого необходимо собрать OpenVINO вместе с модулем openvino\_contrib позволяющим запускать модели на ARM процессорах
- Запустить модели из фреймворка OpenCV на NCS2
- Сконвертировать стороннюю модель повышения разрешения в формат использующийся в OpenVINO, использовать модуль openvino model optimizer



## Заключение

В ходе данной работы были рассмотрены различные платформы, на которых возможен запуск нейронных сетей для повышения разрешения изображений. Были собраны материалы, которые можно использовать в методических целях. Так же были рассмотрены различные фреймворки использующиеся для запуска подобных алгоритмов. Фреймворк OpenCV достаточно удобен для решения широкого спектра задач компьютерного зрения, в него встроена поддержка 4 моделей, которые позволяют решать широкий спектр задач. Однако OpenCV не имеет такой гибкости в добавлении новых моделей как OpenVINO. Был использован NCS2, который позволяет разгрузить процессор, перенести на него вычисления и значительно увеличить скорость вычислений.

## Библиографический список

1. Introduction to Intel OpenVINO Toolkit [электронный ресурс]. - Режим доступа: <https://learnopencv.com/introduction-to-intel-openvino-toolkit/> (дата обращения 01.02.2022)
2. Running OpenVINO Models on Intel Integrated GPU [электронный ресурс]. - Режим доступа: <https://learnopencv.com/running-openvino-models-on-intel-integrated-gpu/> (дата обращения 01.02.2022)
3. Deep Learning based Super Resolution with OpenCV [электронный ресурс]. - Режим доступа: <https://towardsdatascience.com/deep-learning-based-super-resolution-with-opencv-4fd736678066> (дата обращения 01.02.2022)
4. OpenVINO Documentation [электронный ресурс]. - Режим доступа: <https://docs.openvino.ai/latest/index.html> (дата обращения 01.02.2022)

## ОБУЧЕНИЕ МНОГОСЛОЙНОГО ПЕРСЕПТРОНА С ИСПОЛЬЗОВАНИЕМ ГЕНЕТИЧЕСКОГО АЛГОРИТМА

О.В. Судаков

ORCID: 0000-0003-3723-5973 e-mail: [Sudakov.oleg123@yandex.ru](mailto:Sudakov.oleg123@yandex.ru)

Д.В. Дмитриев

ORCID: 0000-0001-9481-0968 e-mail: [dmitdmit@mail.ru](mailto:dmitdmit@mail.ru)

О.Н. Корелин

ORCID: 0000-0002-5928-6045 e-mail: [o\\_korelin@mail.ru](mailto:o_korelin@mail.ru)

Нижегородский государственный технический университет им. П.Е. Алексеева

В докладе рассмотрены вопросы, связанные с обучением искусственного интеллекта при отсутствии обучающих пар и учителя, используя генетический алгоритм. Приведен пример алгоритма с описанием элементов и принципом их работы. Показано, как с помощью полного перебора найти многослойный перцептрон, наиболее адекватно решающий поставленную задачу и почему полный перебор не подходит для современных компьютеров. Проведено исследование с целью нахождения методов увеличения скорости получения оптимального перцептрона при выполнении направленного перебора. Предложены методы, позволяющие сократить время, необходимое на нахождение наилучшей настройки сети, а также повышение точности работы конечного перцептрона. Приведен принцип разбиения нейронной сети на элементы, пригодные для кодирования ген и хромосом генетического алгоритма и рассмотрено несколько принципов мутации и скрещивания.

**Ключевые слова:** нейронные сети, обучение, генетический алгоритм, самообучающиеся алгоритмы, обучение без учителя, обучение баз данных, гибридные системы.

## TRAINING OF A MULTILAYER PERCEPTRON USING A GENETIC ALGORITHM

O.V. Sudakov, D.V. Dmitriev, O.N. Korelin

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

**Purpose:** The article discusses issues related to the training of artificial intelligence in the absence of training pairs and a teacher using a genetic algorithm.

**Design/methodology/approach:** An example of an algorithm with a description of the elements and the principle of their operation is given. It is shown how, using a full search, to find a multilayer perceptron that most adequately solves the problem and why a full search is not suitable for modern computers. A study was conducted to find methods to increase the rate of obtaining the optimal perceptron when performing a directional search.

**Findings:** Methods are proposed to reduce the time required to find the best network setup, as well as to improve the accuracy of the final perceptron. The principle of splitting a neural network into elements suitable for encoding genes and chromosomes of a genetic algorithm is given, and several principles of mutation and crossing are considered.

**Originality/value:** The proposed methods can be used in systems from a genetic algorithm and a perceptron to accelerate learning and increase the accuracy of work. Potentially, it is possible to use these methods in other tasks using genetic algorithms.

**Keywords:** neural networks, learning, genetic algorithm, self-learning algorithms, unsupervised learning, data-free learning, hybrid systems.

## **Введение и постановка задачи**

Сегодня линейные алгоритмы уже не могут удовлетворить рынок услуг и товаров в области решения сложных интеллектуальных задач. Активное развитие интернета как средства не только предоставления информации, но и способа взаимодействия с окружающим человека обществом накладывает новые требования к скорости и точности вычислений.

Нейронные сети являются альтернативным решением, когда предпочтительнее быстро получить хорошее решение, чем точное, но медленно. Интеллектуальные алгоритмы, применяются повсеместно: для обработки различных заявлений, к примеру, в органы власти [1], в бизнесе для рекламы товаров и услуг [2], в автопилотах транспортных средств [3], в поиске информации [4] и даже в играх [5].

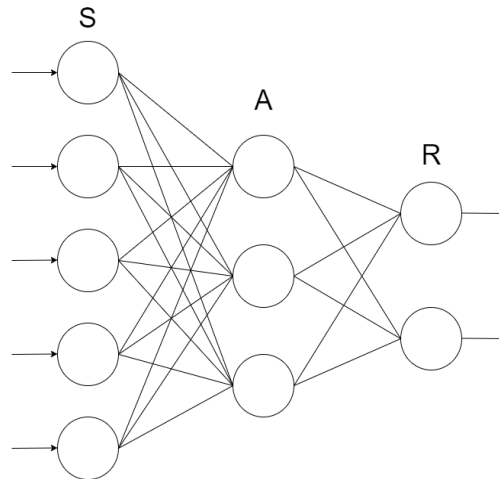
Современные интеллектуальные алгоритмы нуждаются в больших массивах данных, которые помогут обучить нейронную сеть как работать правильно [6]. Но не всегда есть ресурсы для подготовки информации. В этом случае используются специальные интеллектуальные алгоритмы, которые способны сами найти закономерность в исходных данных и обучиться выполнять поставленную задачу. Их так же называют – самообучающиеся интеллектуальные алгоритмы. В данной статье рассмотрено одно из возможных решений, кратко описан принцип работы, а также области применения. Это персептрон, использующий генетический алгоритм для обучения.

Особенностью данного алгоритма является его способность к самостоятельной выработке новых знаний [7]. Таким образом, сеть лучше всего себя проявляет в задачах, где информация для обучения в явном виде отсутствует [8]. Идея заключается в направленном переборе весов. Направление выбирается в соответствии с критерием, характеризующим успешность выполнения задачи [9]. Данный алгоритм не требует данных для обучения, ему необходима симуляция процесса и оценочная характеристика для работы. К сожалению, данный алгоритм хоть и может иметь потенциально достаточно высокую точность при минимальной потребности в данных, имеет серьезные недостатки. Процесс обучения происходит достаточно долго, т.к. множество возможных вариантов слишком велико.

Возможно, данный алгоритм найдет свое применение в будущем, после распространения новых квантовых компьютеров, способных решать задачи перебора намного быстрее современных ЭВМ, хотя сейчас существуют патенты на изобретения, позволяющие ускорить процесс нахождения решения с использованием нынешних технологий [10]. На данный момент подобному виду алгоритмов найдено применение в задачах организации безопасности [11], прогнозировании свойств металлов [12] и т.д.

## **Структура нейронной сети и алгоритмы обучения**

Т.к. далее будут рассматриваться не только цельные системы, но и использующие персептрон, необходимо кратко описать принцип его работы. Он отлично подходит для решения задач классификации и его строение достаточно простое для изучения взаимодействия элементов при обучении.



**Рис. 1. Пример персептрона**

Персептрон имеет несколько составляющих (Рис. 1):

- *S* элементы – элементы, которые принимают входные сигналы.
- *S-A* связи – связи, которые связывают внешний слой входных сигналов и внутренний скрытый слой. *S-A* связи, имеют некоторое значение (вес), на которое умножается значение переходящих по связям данных. Данный вес может быть произвольным и изменяться в некотором диапазоне [13] (в случае персептрона Румельхарта) или иметь значение из множества  $\{-1,0,1\}$  (в случае классического персептрона).
- *A* элементы – это ассоциативные элементы. Данные элементы суммируют все значения, поступившие с входящих связей, а после пропускают сумму через некоторую функцию, обычно сигмовидную. Результат данной функции отправляется в дальнейшие связи, на следующий скрытый или внешний слой.
- *A-R* связи – аналогично *S-A* связывают внешний выходной слой и внутренний скрытый. В отличие от *S-A* связей всегда имеют веса из определенного диапазона, вне зависимости от вида персептрона.
- *R* элементы – образуют слой, который предоставляет результат работы всей сети. Данных элементов, как правило, столько же, сколько и вариантов ответов на поставленную задачу.
- *A-A* связи –используются в многослойных персептронах. Они повторяют собой *A-R* связи, с тем исключением, что связывают не внутренний и внешний слой, а лишь скрытые слои. Их веса берутся из заданного диапазона, не зависимо от вида персептрона.

Имея структуру нейронной сети, необходимо выбрать способ обучения, подходящий целям исследования. Популярный способы обучения – это метод коррекции ошибки, но он требует использования большого объема данных, подготовленных специальным образом и наличия готового верного ответа на любой вариант входа, что противоречит условию поставленной задачи.

Есть метод обучения с учителем, но данный метод также не подходит нам, т.к. необходимо реализовывать систему способную осуществлять те же операции, которым и обучается интеллектуальная система для генерации большого количества обучающих пар, либо искать уже готовые данные. Таким образом данный метод подразумевает выполнение дополнительной работы.

Методы, которые будут рассмотрены далее, используют дополнительные алгоритмы для обработки исходных данных и создания из них обучающих пар или для прямой генерации и перебора различных вариантов персептрона. Комплекс этих алгоритмов образуют самообучающиеся системы, которые так же называют гибридными. Это позволят избавиться от одного из главных недостатков персептрона: необходимости в большом количестве

информации для его обучения и одновременно с этим сохраняет его преимущества, такие как точность и скорость работы.

### Принцип работы генетического алгоритма с персептроном

Рассматривая нейронную сеть, можно увидеть, что на результат ее работы влияют множество параметров, но большинство из них уже predeterminedены и не изменяются непосредственно при обучении. Изменению подвергаются только веса связей, которые соединяют между собой слои нейронной сети. Поскольку все весовые коэффициенты одной сети можно представить в виде нескольких матриц, то при ограничении весов в некотором диапазоне, задача будет иметь конечное число решений.

Возьмем простой пример: пусть дана сеть с 2 входными нейронами и одним выходным (Рис. 2).

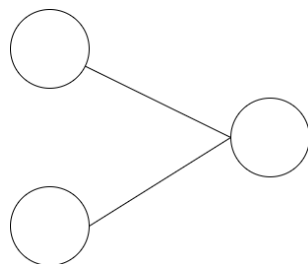


Рис. 2. Пример сети 2-1

Скрытые слои отсутствуют, а весовые коэффициенты могут изменяться в диапазоне [0.00...1.00] с шагом 0.01. Таким образом, используя формулу комбинаторики о количестве размещений:

$$A_n^k = \frac{n!}{(n-k)!} \quad (1)$$

где  $m$ - упорядоченное подмножество элементов,  $n$  – все множество.

Не сложно подсчитать, что количество возможных вариаций данной нейронной сети равно:

$$A_n^k = \frac{n!}{(n-k)!} = 9900 \quad (2)$$

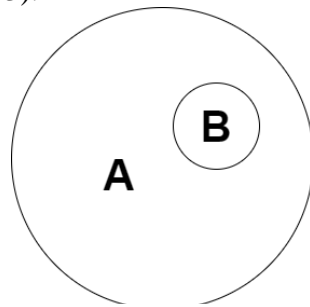
Это означает, что для любой задачи, мы можем опробовать 9 900 подобных нейронных сетей и выбрать ту, которая выполняет задачу лучше остальных. Но проблема заключается в том, что данная нейронная сеть крайне проста и не может решать сложные задачи. Создание же более сложной, с точки зрения архитектуры сети, увеличивает количество вариантов на  $(n-k)! + 1$ . То есть добавление всего одной дополнительной связи в данную сеть увеличит количество возможных вариантов до 970 200 вариантов.

Перебор даже такого количества требует огромного числа ресурсов, а более сложные сети требуют обработки еще большего числа решений. Поскольку данная задача не может быть решена методом простого перебора, было решено использовать генетический алгоритм для создания направленного перебора.

Принцип его работы следующий: несколько нейронных сетей, с различными весами, выполняют одну и ту же задачу. После того, как задача была выполнена, по некоторым критериям отбираются те сети, что выполнили задачу лучше остальных [14]. Веса данных сетей сохраняются, а остальных сетей (показавших результат хуже) корректируется.

Корректировка может производиться различными методами, но результатом должно быть приближение остальных сетей к потенциально лучшему результату [15]. Поскольку набор весов всех сетей изначально определен, то очевидно, что при перестановке весов

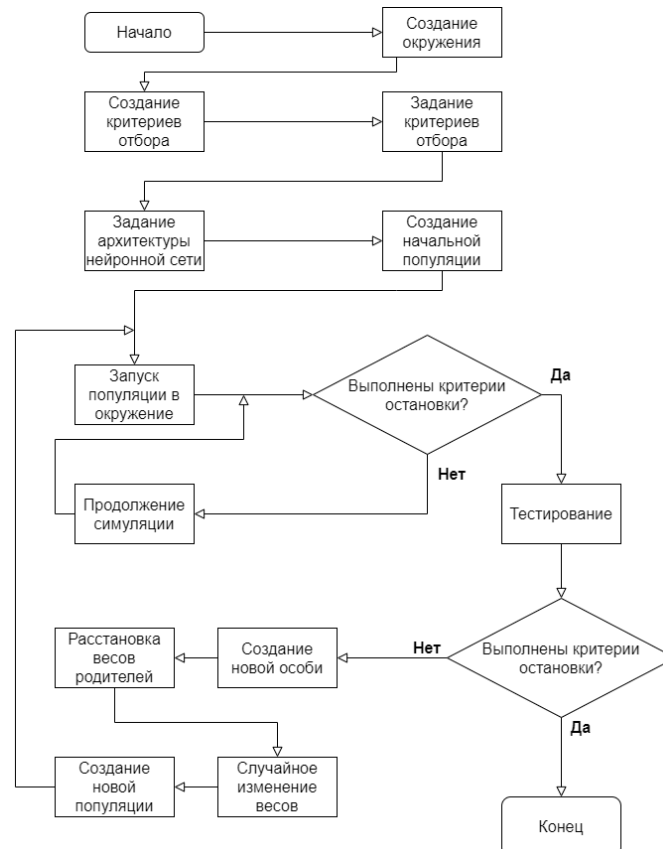
внутри популяции, без изменения самих значений весов, оно составляет лишь некоторое подмножество  $B$  всех весов  $A$  (Рис. 3).



**Рис. 3. Графическое представление рассматриваемых множеств**

Путем скрещивания мы перебираем лишь доступные нам веса во множестве  $B$ . Чтобы выйти за пределы этого множества, необходимо добавить «мутацию».

Мутация – это алгоритм, позволяющий с небольшим шансом изменить некоторые веса в нейронной сети после скрещивания.[16] Тем самым мы можем «перемещать» наше подмножество  $B$  внутри  $A$ . Иллюстрация данного алгоритма можно увидеть на блок-схеме (Рис. 4).



**Рис. 4. Блок-схема генетического алгоритма**

Можно заметить, что размер множества  $B$  можно изменять. Чтобы одновременно проверять большее количество вариантов, достаточно увеличить популяцию одного поколения, что приведет к проверке большего числа различных решений. Но данный алгоритм обладает некоторыми недостатками.[17]

Перебор подразумевает под собой «движение вперед». Иначе говоря, решения, которые уже были проверены, но по каким-либо причинам не подошли, не должны проверяться.

Выполнение данного условия при больших размерах множества  $A$  становится все сложнее, т.к. необходимо большое количество оперативной памяти.

Даже если выделять лишь по 1 Кб на хранение каждого решения, для хранения всех сетей с 3 связями понадобится уже практически 1 Гб. Соответственно более сложные сети заполняют оперативную память гораздо быстрее. По той же причине невозможно создать изначальный вектор или стек и перебирать варианты из него путем отбрасывания неудачных решений.

Для решения данной проблемы перед началом обучения все веса нейронных сетей обнуляются. При подстановке нулей вместо случайных значений сеть выдает одинаковый вариант ответа на любой вход. Тем самым первая популяция проваливается, но после скрещивания происходит мутация. Это активирует некоторые связи, при этом остальные все также не влияют на результат. Тем самым несколько нейронных сетей, с активированными связями, выполняют задачу и те, работа которых повлияла на результат лучше всего, закрепляют веса в следующем поколении. Образуя данным скрещиванием несколько рабочих связей, получается прототип нейронной сети, относительно которой уже начинает происходить корректировка весов, тем самым достигаются наилучшие результаты.

Рассуждая таким образом, можно прийти к выводу, что самым удобным способом организации хромосом является хранение весов между слоями в виде вектора, где каждый вес является геном. Это позволяет проводить мутацию как точно (для каждого гена внутри хромосомы), так и используя хромосомы.

Однако сложность сети накладывает определенные ограничения и сеть с большим количеством связей и слоев обучается значительно дольше, нежели сеть с меньшим количеством, что является вполне ожидаемым т.к. количество весов, значения которых необходимо найти, растет.

### **Проверка эффективности генетического алгоритма**

В качестве экспериментальной апробации совместного использования генетического алгоритма и персептрона использована игра «змейка». Игра выбрана по нескольким причинам:

- Наличие системы рекордов, косвенно показывающей успешность выполнения работы;
- Возможность варьирования размера входных данных;
- Простота разработки;
- Низкие требования к производительности для одной игровой сессии;
- Наличие небольшого элемента случайности;

Рассмотрим особенности взаимодействия нейронной сети с игрой:

- Специальная функция представляет данные для входного слоя в виде вектора, содержащего расстояния до ближайших границ (включая хвост змейки) и расстояние до яблока.
- В качестве выхода использовался бинарный вектор, указывающий на одно из 4 направлений.
- Сеть использует только прямое распространение, т.к. корректировка производится путем замены весов и мутации, а не математически.
- Каждая сеть – отдельная особь в терминах генетического алгоритма. После окончания тестирования поколения, составляется рейтинг лучших особей, в программной реализации представляющий собой вектор из  $N$  особей. Параметр  $N$  можно настраивать с целью определения влияния данного значения на результат.
- С каждой сетью скрещивание производится несколько раз. Число раз определяется рейтингом и является числом, пропорциональным рекорду особи. Рекорд – количество съеденных яблок. При скрещивании каждый вес берется из одной из двух, на данный

момент лучших нейронных сетей. Веса подставляются случайным образом, при сохранении архитектурной позиции относительно предка.

- К сожалению, возможно "Движение назад между поколениями". Причиной этого явления в наличии элемента вероятности в работе системы. Поскольку система генерирует новое яблоко случайным образом, исключается эффект «запоминания». Отрицательной стороной данного решения является возможность потери хороших вариантов в рамках решения одной из подзадач. Тем не менее, эксперимент показал, что лучшим вариантом было именно случайная генерация яблока.

Таблица 1

Результаты экспериментов после 50 поколений

Номер теста	Кол-во особей	Вероятность мутации	Кол-во скрытых слоев	Кол-во нейронов в слоях	Размер вектора, содержащего лучших особей для скрещивания	Макс. рекорд	Ср. рекорд на поколение в конце эксперимента
1	2000	0,05	0	0	2	6	2,150
2	2000	0,20	0	0	2	6	2,089
3	2000	0,05	1	8	2	4	2,046
4	2000	0,20	1	8	2	4	2,019
5	2000	0,05	0	0	10	21	3,281
6	2000	0,05	1	8	10	4	2,008
7	20000	0,05	0	0	2	27	3,014
8	20000	0,20	0	0	2	6	2,029
9	20000	0,05	1	8	2	4	2,051
10	20000	0,20	1	8	2	8	2,038
11	20000	0,05	0	0	10	43	4,206
12	20000	0,05	1	8	10	5	2,022
13	200000	0,05	0	0	2	58	8,192
14	200000	0,20	0	0	2	37	3,184
15	200000	0,05	1	8	2	6	2,053
16	200000	0,20	1	8	2	13	2,069
17	200000	0,05	0	0	10	42	3,648
18	200000	0,05	1	8	10	5	2,050
19	200000	0,10	1	8	1000	11	2,188

По построенной таблице можно сделать следующие выводы (табл. 1):

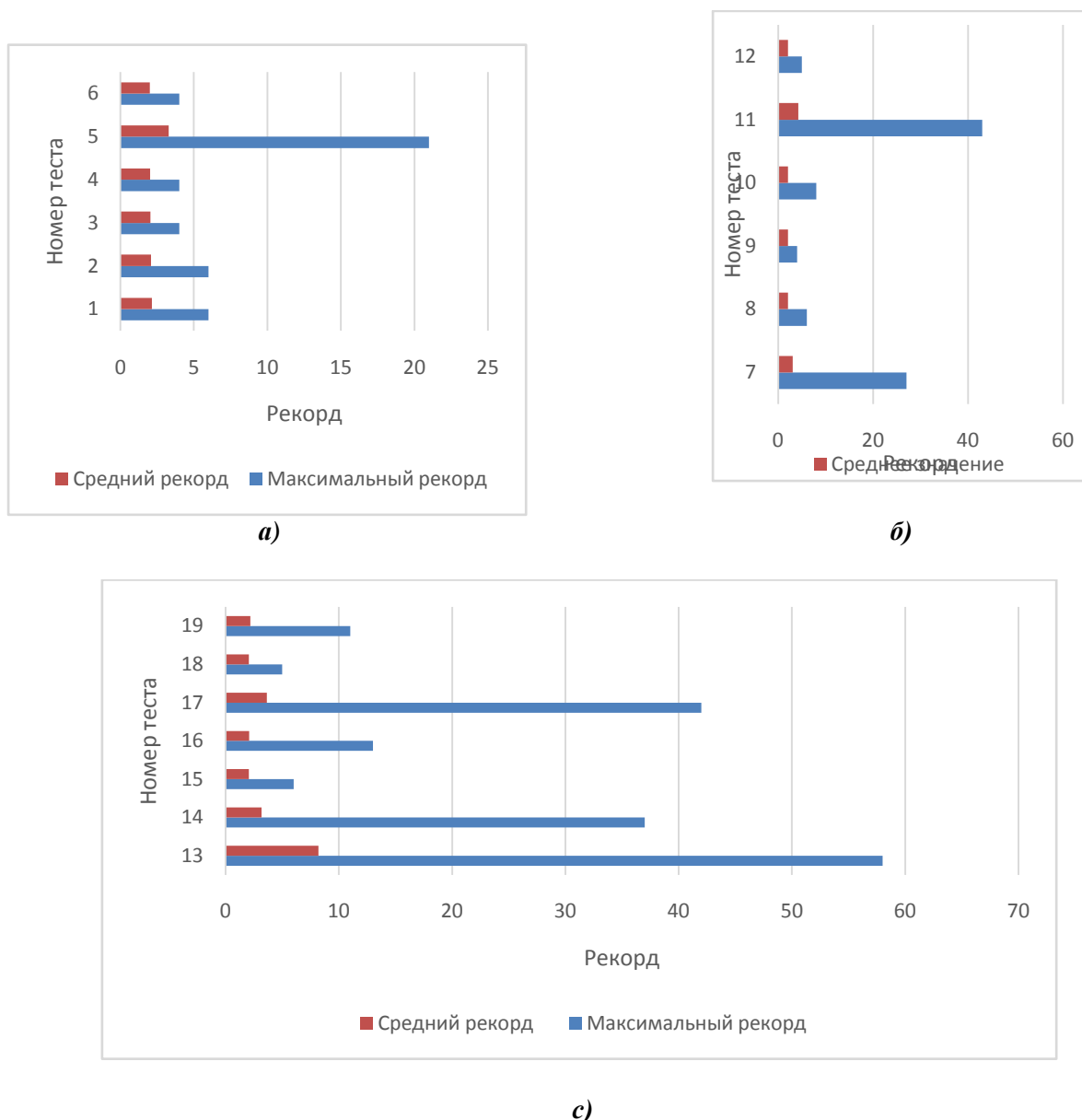
- Алгоритм обучается за меньшее число поколений, когда больше популяция на одно поколение. Математически это подтверждается, если рассматривать популяцию, как исследуемое подмножество  $B$ , являющееся частью всего подмножества  $A$ .
- Добавление скрытого слоя сильно увеличивает количество поколений, необходимое для достижения того же результата без него, но теоретически, увеличивая количество эпох, результат нейронной сети без скрытых слоев для такого алгоритма достижим при любой архитектуре.
- Увеличение шанса мутации может ухудшить обучаемость. Это происходит по причине слишком сильной удаленности от «предка» и потерей накопленного результата.



Однако для нейронной сети со скрытым слоем большое значение вероятности мутации позволяет произвести первичную настройку сети за меньшее число поколений. Следовательно, данный параметр выгодно делать динамическим (Большие значения для грубой настройки и меньшие для тонкой настройки сети).

- Увеличение размера вектора особей для скрещивания хорошо влияет на итоговый результат. Однако стоит учитывать, что слишком маленькое значение дает более точную, но медленную настройку сети, в то время как слишком большое может быстро уменьшить точность персептрона.

Построим графики для сравнения результатов. Отметим среднее значение рекорда для поколения и максимальный достигнутый рекорд (Рис 5 (а-с)).



**Рис. 5** Графики результатов

По данным графикам видно, что большая мутация чаще приводит к снижению показателей. Добавление скрытого слоя увеличивает количество возможных вариантов и замедляет обучение. Скрещивание 10 особей, вместо двух, дает возможность повысить среднее значение, тем самым лучше сохраняя полезные гены.

## Заключение

Мы рассмотрели принцип корректировки весов персептрона методом направленного перебора. Для выбора направления применялся генетический алгоритм. Взаимодействие двух алгоритмов потенциально дает хорошую точность, а также не требует заранее подготовленных данных, что снижает затраты в человеко-часах, необходимых для реализации программного обеспечения, но временные затраты на обучение могут оказаться очень высоки. Данный алгоритм был протестирован на игре «Змейка», т.к. данная программа лучше всего позволяла провести оценку алгоритма, за счет несложных правил и системы рекордов. По результатам тестирования преимущества и недостатки были подтверждены.

Тенденция развития нейронных сетей активно увеличивает возможные области применения и позволяет развивать рассмотренный алгоритм дальше с целью снижения временной нагрузки на программиста и увеличения точности работы программ.

## Список литературы

1. **Бахтаирова, Е. А.** Нейронные сети и искусственный интеллект в работе с обращениями граждан в органы власти / Е. А. Бахтаирова, М. В. Вихорева // Развитие российского общества: вызовы современности: Материалы национальной научно-практической конференции с международным участием, посвященной 90-летию Байкальского государственного университета, Иркутск, 15–16 октября 2020 года. – Иркутск: Байкальский государственный университет, 2021. – С. 120-125.
2. **Кораблев, А. Ю.** Машинное обучение в бизнесе / А. Ю. Кораблев, Р. Б. Булатов // Азимут научных исследований: экономика и управление. – 2018. – Т. 7. – № 2(23). – С. 68-72.
3. **Климин, А. А.** Использование искусственного интеллекта в автомобилях / А. А. Климин, С. С. Гилева, Н. Е. Отекина // Мир Инноваций. – 2020. – № 4. – С. 48-51.
4. **Коваленко, Т. А.** Искусственный интеллект в современном мире / Т. А. Коваленко, Р. В. Никифоров // Научное пространство: актуальные вопросы теории и практики : материалы II Всероссийской научно-практической конференции, Чебоксары, 24 сентября 2021 года. – Чебоксары: Негосударственное образовательное частное учреждение дополнительного профессионального образования "Экспертно-методический центр", 2021. – С. 11-15.
5. **Кононенко, Д. В.** Развитие искусственного интеллекта в играх и применение в других отраслях / Д. В. Кононенко, А. М. Хахина // Информационные технологии: прошлое, настоящее, будущее : сборник статей по материалам межинститутской студенческой научно-практической конференции, Санкт-Петербург, 20 мая 2021 года. – Санкт-Петербург: Санкт-Петербургский государственный лесотехнический университет имени С.М. Кирова, 2021. – С. 172-177.
6. **Короткова, Е. В.** Машинное обучение и нейронные сети. Важный шаг в развитии человечества / Е. В. Короткова // Современные вопросы естествознания и экономики : Сборник трудов III Международной научно-практической конференции, Прокопьевск, 18 марта 2021 года. – Прокопьевск: Филиал Федерального государственного бюджетного образовательного учреждения высшего профессионального образования "Кузбасский государственный технический университет имени Т. Ф. Горбачева" в г. Прокопьевске, 2021. – С. 240-242.
7. **Кондратьев, Т. Н.** Эволюционные вычисления: нейронные сети и генетические алгоритмы / Т. Н. Кондратьев // Международная конференция по мягким вычислениям и измерениям. – 2019. – Т. 1. – С. 418-421.
8. **Загинайло, М. В.** Обучение искусственных нейронных сетей с помощью математического аппарата генетических алгоритмов / М. В. Загинайло // Инновации. Наука. Образование. – 2020. – № 12. – С. 384-389.
9. **Д. А. Петросов.** Применение искусственных нейронных сетей в задачах управления генетическим алгоритмом /Р. А. Ващенко, А. А. Степовой// Моделирование, оптимизация

- и информационные технологии. – 2019. – Т. 7. – № 4(27). – С. 10-11. – DOI 10.26102/2310-6018/2019.27.4.016.
10. Патент № 2602973 С1 Российская Федерация, МПК G06N 3/08, G06N 3/12, G06F 15/18. Контроллер управления обучением нейронной сети с генетическим алгоритмом : № 2015141707/08 : заявл. 30.09.2015 :опубл. 20.11.2016 / В. И. Ключко, Е. А. Шумков ; заявитель Федеральное государственное бюджетное образовательное учреждение высшего образования "Кубанский государственный технологический университет" (ФГБОУ ВО "КубГТУ").
  11. **Амосов, О. С.** Использование генетического алгоритма для настройки глубоких нейронных сетей в задачах классификации компьютерных атак / О. С. Амосов, С. Г. Амосова, Д. С. Магола // Информатика и системы управления. – 2020. – № 4(66). – С. 104-117. – DOI 10.22250/isu.2020.66.104-117.
  12. **Нургаянова, О. С.** Обучение нейронной сети для прогнозирования свойств никелевых сплавов на основе генетического алгоритма / О. С. Нургаянова, Н. И. Юсупова // Труды Института системного анализа Российской академии наук. – 2019. – Т. 69. – № 4. – С. 22-28. – DOI 10.14357/20790279190403.
  13. **Т. Рашид** Создаем нейронную сеть // Лондон: Диалектика-Вильямс– 2017.
  14. **Марков А. Д.** Генетические алгоритмы. Сравнение способов выбора родителей для скрещивания // StudNet. – 2020. – №9.
  15. **Свечников Д. А.** Генетические Алгоритмы / Кузьмин Н. А., Мтирелишвили М. Д., Забелин Р. Р., Лузгарев В. Ю., Гнусарев Я. Ю., Панченко А. С., Чембулатов А. Б. // Наука, образование и культура. – 2020. – №7 (51).
  16. **Ведерникова О.Г.** Динамические нормы для случайных и направленных мутаций в генетическом алгоритме /Игнатьева О.В. // ИВД.– 2019.– №5 (56).
  17. **Есер F.** et al. Training multilayer perceptron with genetic algorithms and particle swarm optimization for modeling stock price index prediction //Entropy. – 2020. – Т. 22. – №. 11. – С. 1239.
  18. **Wang H.**Genetic algorithm hybridized with multilayer perceptron to have an economical slope stability design / Moayedi H., KokFoong L. //Engineering with Computers. – 2021. – Т. 37. – №. 4. – С. 3067-3078.

**РАЗРАБОТКА И ИССЛЕДОВАНИЕ АЛГОРИТМОВ СОПРОВОЖДЕНИЯ ОБЪЕКТОВ  
НА ВИДЕОПОТОКЕ****М.Б. Багиров**ORCID: 0000-0003-1656-0849 e-mail: [bagirov\\_mirabbas@mail.ru](mailto:bagirov_mirabbas@mail.ru)

Нижегородский государственный технический университет им. П.Е. Алексеева

В докладе представлены результаты разработки интеллектуальной системы, применяемой для анализа транспортных средств на видеопотоке – детектирование, классификация, отслеживание транспортных средств. Для тестирования работы системы использовался StanfordCarDatasetс добавлением собственного датасета, включающего в себя 6000 изображений. Разработанная система может использоваться в интеллектуальных транспортных системах (ИТС). Решение имеет микросервисную архитектуру с легко обновляемыми и заменяемыми модулями в соответствии с требованиями функциональности системы. Данная задача не тривиальна, так как требуется учет большого количества факторов (положение камеры, условия освещенности, перекрытие объектов, разрешения изображений в видеопотоке и т.д.), что делает задачу актуальной.

**Ключевые слова:** компьютерное зрение, анализ изображений, нейросети, точность распознавания, классификация транспортных средств

**DEVELOPMENT AND RESEARCH OF ALGORITHMS FOR TRACKING OBJECTS  
ON A VIDEO STREAM****M.B. Bagirov**

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

**Purpose:** to study and implement computer vision algorithms for analyzing vehicles in a video stream.

**Design / Methodology / Approach:** research was conducted and algorithms were developed to detect, classify, and track vehicles in a video stream. In studying the classification algorithms, the ResNet50 and VGG16 models were compared using error (loss) and accuracy metrics. As a result of the comparison, the ResNet50 model was selected to solve the classification problem, of which a sub-model was also selected to use the penultimate layer as the output layer. It was assumed that the vector obtained at the output contains enough representative features of the car in the image, based on which new classifiers can be created using clustering algorithms that do not require training with a teacher without retraining the original neural network. In this way, the tasks of classifying vehicles in the video stream according to the desired attribute could be successfully solved. An algorithm based on the Siamese triplet network (ReID) using the methods of SORT was used to track vehicles in the video stream. The results of the research and image processing are presented in the form of graphs, tables, and labeled results on images. For training the models, the Stanford Car Dataset was used, which contains 16,185 images classified into 196 classes, enriched with its dataset of 6,000 images whose objects characterize features such as make, model, year of manufacture, body type, color, and shooting angle. The markup was performed using the Computer Vision Annotation Tool (CVAT). In training, Adam optimizer was used as optimizer and categorical cross entropy as loss function.

**Findings:** The article presents the results of data mining on a video stream for vehicle detection, classification and tracking to further adapt them to specific tasks.

**Research limitations/implications:** The Stanford Car Dataset was used for this work, which was enriched with 6000 images from our own dataset. The photos were taken at different times of the day and under different lighting conditions to increase the stability of the developed solutions. The developed system has a microservice architecture that makes it easy to update, replace and add new modules depending on the functional requirements of the system.

**Originality / Value:** Currently, the analysis of the video stream and the monitoring of the vehicles are performed manually by the staff. Automating this process will not only reduce the workload and cost for the operator, but also increase the speed of decision making and the accuracy of the reports provided. The developed system can be used in intelligent transportation systems (ITS). The solution has a microservice architecture with modules that can be easily updated and replaced depending on the functional requirements of the customer.

**Keywords:** computer vision, image analysis, neural networks, recognition accuracy, vehicle classification

## **Введение и постановка задачи**

В настоящее время большой интерес представляет использование систем «умной» видеоаналитики, разработанных на основе технологий компьютерного зрения. Данные технологии актуальны в различных областях экономики, госсектора, бизнеса, образования. При этом в качестве отраслевых драйверов роста решений с использованием видеоаналитики в настоящее время выступают транспорт, здравоохранение, производство и ритейл.

Существуют различные закрытые продукты компаний, которые используются в комплексных решениях безопасности транспортных систем, включающих дорожное видеонаблюдение с фиксацией нарушений, контроль за ситуацией на парковках транспортных средств, распознавание и сопровождение транспортных объектов, контроль въезда/выезда транспорта с охраняемых территорий. Готовые решения продуктовых компаний, использующие готовые алгоритмы различных библиотек и пакетов, встраивание которых делает продукт не масштабируемым, не всегда удовлетворяют своим качеством заказчиков, поскольку часть из них невозможно адаптировать под конкретные задачи заказчика, они дорогие, и в меняющихся реалиях достаточно быстро утрачивают актуальность. К достаточно большому количеству готовых и широко используемых решений можно отнести сбор статистики дорожного движения, распознавания госномеров транспортных средств, детектирование скорости и остановок транспортных средств. Меньше решений в области автоматизированной фиксации нарушений, угроз, создаваемых транспортными средствами, например, выявление ДТП, классификация опасного вождения и выявление происшествий с участием пешеходов.

При этом архитектура этих решений не позволяет доработать и встроить в них новые уникальные модули для решения возникающих проблем. Необходимо разрабатывать и внедрять инновационные решения для покрытия все новых требований в области распознавания и детектирования объектов в больших потоках транспортных средств, где сопровождение объекта является сложной задачей из-за постоянных перекрытий, перестроений, смены ракурса транспортного средства. Большое количество конкурсов и хакатонов для IT-специалистов в сфере разработки новых решений для обеспечения информационной безопасности транспортных систем подчеркивает актуальность научных исследований и разработки новых решений в сфере видеоаналитики, внедрения новых сервисов для интеллектуализации транспортных систем.

В настоящее время применение искусственного интеллекта актуально в системах мониторинга транспортных средств для реализации функционала распознавания типа, марки, модели и цвета транспортных средств. Существуют решения определения такого рода параметров транспортных средств по фотоизображениям с использованием датасетов *SortmCars* и *Cars*. Решение задач анализа видеопотоков для распознавания и классификации по типам, маркам и моделям транспорта представляет в настоящее время значительные трудности, связанные в том числе с отсутствием датасетов.

Давно и успешно при определенных условиях применяются системы распознавания госномеров транспортных средств. Однако остается не решенной задача распознавания госномеров от следящих камер при увеличении диапазона параметров перспективных искажений, возникающих в результате наклона и поворота номерной пластины относительно оси камеры. Помимо требований стандарта шрифта и размера, контрастности между символами номера и фоном, существуют достаточно жесткие ограничения на распознавание номеров при отклонениях плоскости номерной пластины – угол от оси камеры не должен превышать 300, наклон по горизонтали не более 150. В данной работе представлены исследования в этой области и предложены новые решения, позволяющие улучшить качество и расширить функционал применения видеоаналитики для обработки транспортных потоков.

Исследование решений в области анализа видеопотока движения транспортных средств показал, что большинство из них, при незначительных корректировках, могут быть применены для решения задач анализа видеопотоков в других сферах деятельности. Это говорит о том, что они не обладают свойством уникальности, а используют известные модели и алгоритмы, настройка которых часто не дает требуемого качества решений. Следует отметить, что мониторинг дорожного движения представляет сложный непрерывный процесс сбора данных и их анализа в режиме реального времени – наблюдение, выделение определенных признаков и свойств, распознавание и отслеживание движения различных сущностей, а также детализация объектов. Большое значение здесь имеют характеристики как аппаратных средств, так и применяемые алгоритмы, учитывающие разные факторы – положение камеры, условия освещенности, качество изображения на видеопотоке. При этом необходима интеграции аналитических инструментов, реализованных в IP-камерах, с разрабатываемым программным обеспечением. Кроме того, разрабатываемый интерфейс должен быть интуитивно понятен для пользователя.

При разработке алгоритмов определения параметров – типа, марки и модели транспортных средств на данный момент существует несколько проблем. Исследования показали, что при использовании в решениях алгоритма центроидов существуют перекрытия, ведущие к потери объекта, существует алгоритм на основе векторного анализа, в основе которого лежит выделение вектора признаков изображения автомобиля для реализации которого используется классификационная сверточная нейронной сети и дальнейшее сравнение его с образцовым вектором для исследования схожести. Данный алгоритм ведет к большому проценту ошибки в случае определенного сходства находящихся рядом транспортных средств.

Разработка системы интеллектуального мониторинга за движением транспортных средств включает в себя следующие задачи:

- детектирования транспортных средств на видеопотоке;
- классификации транспорта по типу, марке и модели;
- отслеживания (сопровождения) транспортного средства на видеопотоке;
- оптимизация моделей нейронных сетей для решения поставленных задач анализа;
- визуализации полученных данных.

### **Детектирование транспортного средства на видеопотоке**

Задача детектирования объектов актуальна и решается несколько десятилетий, за которые было разработано множество алгоритмов, которые основаны на выявлении признаков, присущих объекту детектирования. Их можно разбить на два больших класса: алгоритмы, основанные на классических подходах и алгоритмы глубокого обучения с использованием нейронных сетей.

Среди подходов можно выделить детектирование с использованием гистограмм направленных градиентов, метод Виолы-Джонса, а также алгоритмы, основанные на детектировании объектов по контуру. Первые два являются достаточно универсальными и требуют лишь переобучения на нужные сущности, в то время как третий является, по сути, целым классом различных алгоритмов, которые необходимо разрабатывать под каждую конкретную задачу.

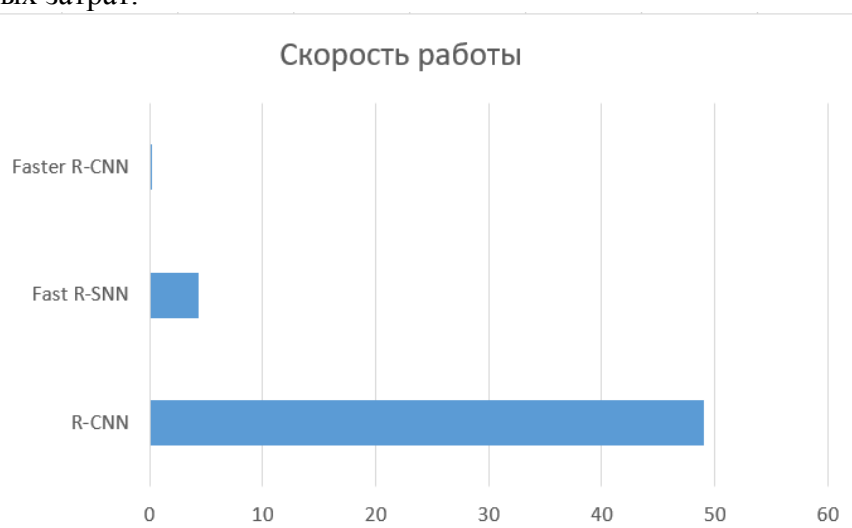
В данной задаче будут использоваться алгоритмы глубокого обучения. К наиболее популярным алгоритмам, используемым в задачах анализа видеопотока в интеллектуальных транспортных системах относятся:

- Сеть R-CNN (Region-based Convolutional Network). Анализ показывает, что R-CNN не может быть использована в режиме реального времени, так-как является энергозатратным алгоритм, требующим много времени на обучение, так-как применяется метод selective search[8].

- Сеть Fast R-CNN. Анализ показывает[9], что медленный и сложный процесс тренировки сети не является большим недостатком, в то время как временные затраты на этапе детектирования сильно снижают практическую пользу от применения сети.

Fast R-CNN быстрее, чем R-CNN, поскольку операция свертки выполняется только один раз для каждого изображения, и из него генерируется карта объектов.

- Сеть Faster R-CNN. Алгоритмы R-CNN и Fast R-CNN используют выборочный поиск, который замедляет и удлиняет процесс, влияющий на производительность сети. Faster R-CNN лишена данного недостатка и использует свой алгоритм локализации Region Proporsal Networks[10], что позволяет добиться высокой производительности и низких временных затрат.



**Рис. 1. Скорость работы алгоритмов R-CNN, Fast R-CNN, Faster R-CNN**

Из приведенного графика, представленного на Рис. 1, видно, что скорость Faster R-CNN значительно выше, чем R-CNN и Fast R-CNN. Таким образом, Faster R-CNN может использоваться для обнаружения объектов в реальном времени.

- YOLO (You Only Look Once) [1]. Анализ показывает, что скорость предложенных выше сетей не сильно высока для детектирования объектов, так как они используют области для локализации объекта в изображении. YOLO отличается от других сетей за счет иного подхода к проблеме детектирования и имеет два больших преимущества:
  - Позволяет производить одновременное детектирование нескольких разных типов сущностей, т. е. фактически она способна детектировать и одновременно распознавать объекты;
  - В отличие от большинства других алгоритмов детектирования, YOLO не использует пирамиду изображений и проходит плавающим окном по изображению лишь единожды, что положительно сказывается на времени работы алгоритма.

В данной работе рассматривается пятая версия сети, включающая в себя 4 модели разных размеров: YOLOv5s, YOLOv5m, YOLOv5l, YOLOv5x. Модели YOLOv5s, YOLOv5m и YOLOv5x не отличаются слоями. Разница между четырьмя версиями сети YOLOv5 заключается в компромиссе модели и времени вывода. В таблице 1 приведено сравнение моделей YOLOv5.

Таблица 1

Сравнение моделей YOLOv5

Model	size (pixel)	mAPval 0.5:0.5	mAPtest 0.5:0.5	mAPval 0.5	Speed V100(ms)	params (M)	Flops 640 (B)
YOLOv5s	640	36.7	36.7	55.4	2.0	7.3	17.0
YOLOv5m	640	44.5	44.5	63.1	2.7	21.4	51.3
YOLOv5l	640	48.2	48.2	66.9	3.8	47.0	115.4
YOLOv5x	640	50.4	50.4	68.8	6.1	87.7	218.8
YOLOv5s6	1280	43.3	43.3	61.9	4.3	12.7	17.4
YOLOv5m6	1280	50.5	50.5	68.7	8.4	35.9	52.4
YOLOv5l6	1280	53.4	53.4	71.1	12.3	77.2	117.7
YOLOv5x6	1280	54.4	54.4	72.0	22.4	141.8	222.9
YOLOv5x6 TTA	1280	55.0	55.0	72.0	70.8	-	-

Для обучения использовалась модель YOLOv5m. Демонстрация обученной модели представлена на Рис. 2.

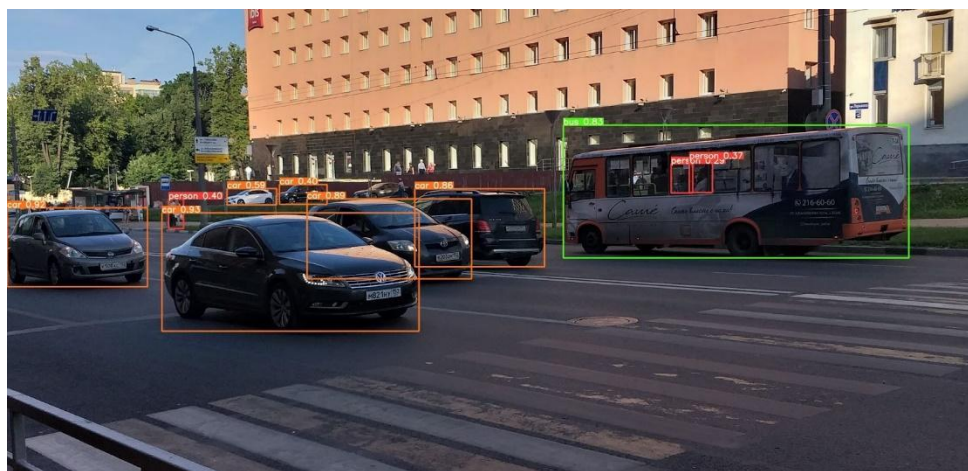


Рис. 2. Демонстрация работы YOLOv5m

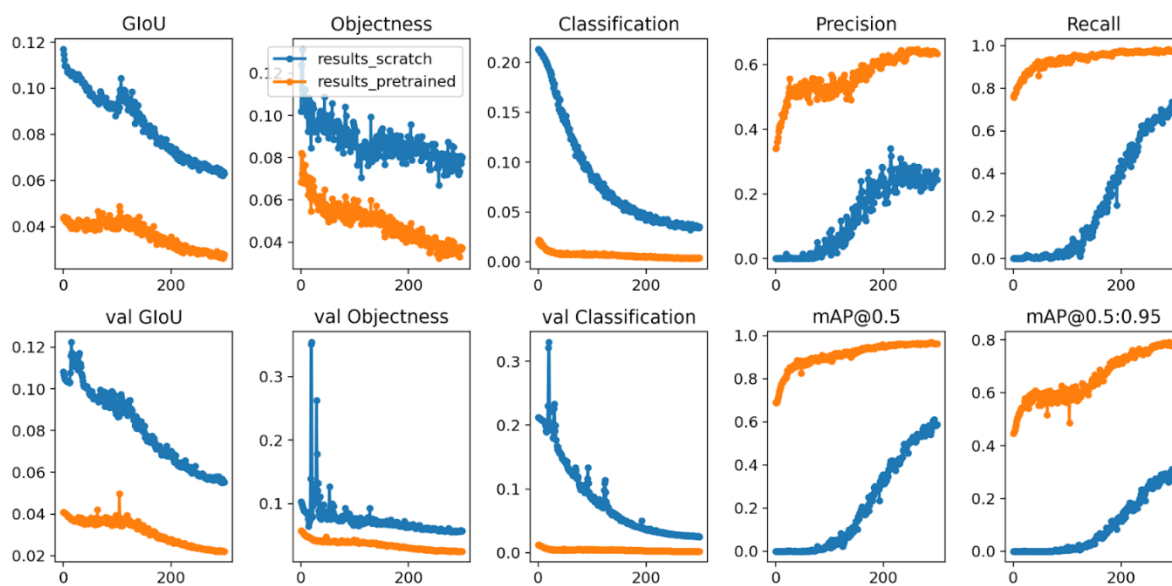


Рис. 3. Динамика показателей качества в процессе обучения YOLOv5m (синяя линия)



### Выделение признаков и распознавание транспортного средства

В результате исследований в области выделения признаков и распознавания параметров транспортных средств было проведено сравнение сетей VGG, AlexNet, BN-inception и ResNet, на датасете ImageNet. В результате для классификации транспортных средств по типу, модели и марке была выбрана и исследована модель нейронной сети ResNet, показавшая минимальный процент ошибок на тестовой выборке. Данная модель была дообучена на собранном наборе данных, состоящем из Stanford Cars Dataset [2], а также собранных в полуавтоматическом режиме изображений с сервисов Avito и CarPrice. Собранный набор содержит 6000 фотографий, на каждой из которых выделены следующие признаки: тип кузова, марка, модель, год выпуска, цвет. Все собранные фотографии разнообразны - сделаны в разное время суток, с разных ракурсов и при разных условиях освещенности, что позволило значительно повысить устойчивость алгоритма к изменению внешних условий съемки. Обучение нейросети проводилось с использованием видеокарт, на GPU-сервере, что позволило значительно сократить время тренировки – если на компьютере архитектуры Intel(R) Core(TM) i5-4690 CPU время обучения сети составило 20 часов, то использование графического процессора GeForce GTX 1080 Ti позволило обучить сеть за 3 часа, значение точности составило 96%. Для увеличения вычислительной мощности сервера дополнительно использовалось программное средство CUDA и соответствующая библиотека cuDNN. Расчеты, связанные с обучением нейросети, представляют собой матрицы большого размера, требующие высокой производительности, которая достигается с помощью CUDA-архитектуры.

Тестирование скорости работы разработанного модуля проводилось на ЭВМ, оснащенной 4-ядерным процессором Intel(R) Core(TM) i5-4690 и видеокартой GeForce GTX 1080 Ti.

Большую роль в высокой скорости работы CUDA играет версия драйверов. В таблице 2 приведены средние замеры времени в миллисекундах, которое было затрачено на обучение одного скрытого слоя за одну эпоху, но с разной версией драйверов. Повышение производительности при использовании драйвера версии 3.0 по сравнению с версией 2.2 составляет приблизительно 9%.

Таблица 2

Время обучения одного слоя нейронной сети на видеокартах с разными версиями драйверов

Номер слоя	Версия драйверов		Выигрыш во времени, %
	2.3, мс	3.0, мс	
32	141	112	21,6
128	438	393	11,0
224	1515	1391	8,2
320	3813	3531	7,4
416	7875	7250	7,9
512	14094	12969	8,0

Библиотека cuDNN обеспечивает высокую производительность параллельных вычислений. Её использование обусловлено следующими преимуществами:

- Контекстные API, которые обеспечивают простую многопоточность;
- Оптимизация под новейшие архитектуры NVIDIA GPU;
- Операции по свертке функций, включая кросскорреляцию, которые предназначены для конволюционных нейронных сетей.

Применяя CUDA-технологии, можно ускорить обучение нейросети по сравнению с CPU до 5 раз на GTX 1080 Ti. Время обучения состоит из времени работы с файлами и

минимального перемещения данных между host и device. На видеокарте Nvidia GeForce GTX 1080 Ti при увеличении количества нейронов в слоях от 128 преимущество во времени обучения начинает снижаться, а слои с количеством нейронов больше 512 обучаются медленнее, чем на CPU. Следовательно, в двухслойных нейронных сетях с равным количеством нейронов во входном и выходном слоях применение CUDA- технологии на данной видеокарте является более эффективным при обучении небольших нейросетей.

Таким образом, применение CUDA и библиотеки cuDNN целесообразно при обучении нейронных сетей.

### Нейронная сеть ResNet

Для распознавания признаков была обучена нейронная сеть на основе ResNet 50 [3]. На вход разработанной сети подавались цветные изображения размером  $224 \times 224$  пикселя. В качестве выходных используется два слоя:

- предпоследний слой, содержащий 512 нейронов;
- последний слой, количество нейронов которого равно количеству возможных классов (комбинаций всех признаков в обучающем наборе данных).

Для начальной инициализации ResNet были взяты веса из предобученной модели на наборе данных ImageNet.

Layer (type)	Output Shape	Param #
resnet50 (Functional)	(None, 7, 7, 2048)	23587712
global_average_pooling2d (Gl	(None, 2048)	0
dense (Dense)	(None, 256)	524544
dropout (Dropout)	(None, 256)	0
batch_normalization (BatchNo	(None, 256)	1024
dense_1 (Dense)	(None, 512)	131584
dense_2 (Dense)	(None, 196)	100548
Total params: 24,345,412		
Trainable params: 24,291,780		
Non-trainable params: 53,632		

Рис. 4. Архитектура нейронной сети для распознавания признаков автомобиля

Обучение происходило на протяжении 50 эпох, пакетами по 16 изображений. В качестве функции потерь использовалась категориальная кросс энтропия, а в качестве оптимизатора - Adam optimizer [4]. Оптимизация велась относительно последнего выходного слоя.

После обучение сети была выделена подмодель, в качестве выходного которой использовался предпоследний слой, выдающий на выходе 512-мерный вектор. Сделано предположение, что данный вектор содержит достаточно репрезентативные признаки автомобиля на изображении на основании того, что последний слой сети может производить классификацию. При этом невозможно однозначно интерпретировать каждое значение признаков, однако, на основании этих признаков можно использовать алгоритмы кластеризации для создания новых классификаторов, не переобучая при этом исходную нейронную сеть.

С помощью исходной модели можно классифицировать автомобили на классы из обучающего набора данных без необходимости прибегать к дополнительным алгоритмам машинного обучения, именно поэтому она также использовалась в работе. Пример классификации автомобиля из тестовой выборки представлен на рисунке 5.



**Рис. 5. Определение марки, модели, типа автомобиля**

Для отличия автомобилей между собой на видеопотоках необходимо сравнение расстояния между векторно-признаковым представлением каждого автомобиля между кадрами на видеопотоке с двух разных камер. Так как обучающий набор данных имел фотографии с разным освещением и под разным углом, то работа алгоритма должна производиться устойчиво. В качестве метрики близости используем евклидово расстояние. На рисунке представлены примеры отделения изображений, на которых изображены разные автомобили, снятые под разными ракурсами и, наоборот, связывание изображений одного и того же автомобиля, используя расстояние в пространстве признаков выделенных разработанной нами сверточной нейронной сетью.

Для задачи классификации автомобилей по признаку используются алгоритмы кластеризации, для чего берется набор данных, содержащий изображения ТС разных классов. Для каждого выделяется вектор признаков путем представления изображения в виде точки в 512-мерной пространстве и разбиения получившихся изображений на нужное количество кластеров. Получившиеся кадры должны быть близки к разбиению на классы за счет хорошей репрезентации целевых классов (при этом сами классы были зависимы от ряда внешних признаков автомобиля). Таким образом, мы получили классификатор без обучения с учителем.

Проверим работу системы на примере изображений. Векторизуем каждое изображение и алгоритмом K-means [5] разобьем их на два кластера. Убедимся, что изображения с разными цветами автомобилей попали в разные кластеры.



```
array([1, 1, 1, 1, 0, 0, 0, 0], dtype=int32)
```

**Рис. 6. Разбиение изображения с разными цветами автомобилей на кластеры**

Таким образом, мы без учителя, на основе алгоритма выделения признаков и Kmeans алгоритма кластеризации обучили модель для классификации цвета автомобиля. На изображении можно убедиться, что модель работает корректно и верно предсказывает истинный класс. При этом нейронная сеть изначально не обучалась на автомобилях данной модели.



`array([1], dtype=int32)`

**Рис. 7. Классификация цвета автомобиля**

Таким образом, разработанное нами решение может быть обучено на размеченном наборе данных для классификации автомобилей по различным признакам, а также может успешно использоваться для сравнения внешнего сходства автомобилей на разных изображениях.

### **Отслеживание ТС на видеопотоке**

Задача отслеживания объектов между кадрами является одной из самых сложных в области компьютерного зрения. И существует множество алгоритмов ее решения, однако большинство из них крайне зависимы от внешних условий и требует переобучения для каждого конкретного случая применения.

Простейшим алгоритмом трекинга является трекинг по расстоянию между центроидами. Он связывает объект на новом кадре с объектом на предыдущем кадре по минимальному расстоянию между их центроидами. Но такой алгоритм не устойчив в случае, когда в кадре два объекта располагаются близко друг к другу, при этом вероятность ошибки растет, если обрабатывается не каждый кадр в видеопотоке.

Более сложные алгоритмы кроме расстояния между центроидами учитывают также размеры объекта, скорость движения между кадрами в течении небольшого промежутка времени и другие физические характеристики, в зависимости от того, какой именно объект отслеживается. Использование лишь некоторых параметров обычно не обеспечивает высокую точность работы трекера, а использование множества параметров приводит к низкой скорости работы.

Третий тип алгоритмов трекинга - это оптические трекаеры, которые делятся на трекаеры на основе графов и на основе корреляционных фильтров.

Трекаеры на основе графов строят граф, каждый уровень которого соответствует кадру из обучающего видео, узлы соответствуют объектам на видео, а ребра - взаимосвязи между одинаковыми объектами на соседних кадрах. Изначально каждый из узлов одного слоя соединен со всеми узлами следующего слоя, но в процессе работы путем сравнения близости векторов признаков объектов в узлах, которые данное ребро соединяет, выбираются «активные» ребра, которые соответствуют траектории движения каждого из объектов. Таким образом предсказывается наиболее вероятная траектория для каждого из объектов на некотором количестве кадров. Данный подход показывает хорошую точность работы, однако он не применим в системах, скорость работы которых должна быть близка к реальному времени (так как обрабатывается не два соседних кадра, а целый пакет за некоторый промежуток времени).

Трекаеры на основе корреляционных фильтров являются одними из самых популярных на сегодняшний день. Они обеспечивают высокую скорость, но при этом их точность близка к трекаерам на основе графов. Самым популярным на данный момент трекаером из этой категории является SORT, основанный на фильтре Калмана [6] и венгерском алгоритме.

В данной работе был выбран алгоритм DeepSort [7] в качестве рассматриваемого и целевого для разрабатываемого программного решения, так как он обеспечивает высокую точность и скорость работы. Он основан на близости по расстоянию Махаланобиса, корреляционном фильтре Калмана и сиамской сверточной нейронной сети ReID [8], для реидентификации объекта по его внешнему виду. Последний алгоритм требует обучения для работы с автомобилями.

В качестве оптимизатора использовался Adam Optimizer. Для обучения сиамской сети использовался тот же набор данных что и для обучения ResNet. Обучение происходило на протяжении 50 эпох. В результате была обучена сиамская нейронная сеть, которая в совокупности с общими методами алгоритма SORT помогает отслеживать движение автомобилей на видеопотоке.

### **Заключение**

Таким образом, в ходе исследования были достигнуты следующие результаты:

Изучены алгоритмы детектирования сущностей: метод Виолы - Джонса, нейронная сеть YOLO v5, а также кратко ряд других алгоритмов, с обоснованием почему именно они были выбраны для дальнейшего исследования

Проведен сравнительный анализ, с учетом показателей качества и производительности, в задаче детектирования автомобилей. В результате были выявлены слабые и сильные стороны каждого из подходов, в частности самым эффективным показал себя алгоритм на основе нейронной сети YOLO v5.

Разработаны нейросетевые алгоритмы на основе архитектуры ResNet для выделения признаков автомобилей. А также был адаптирован алгоритм DeepSort для отслеживания автомобилей между кадрами.

Создан прототип веб-сервиса для осуществления мониторинга за транспортными средствами на основе клиент-серверной архитектуры и удобного адаптивного графического веб-интерфейса, позволяющего наблюдать за процессом мониторинга, а также смотреть информацию в таблицах базы данных.

Произведено тестирование в нескольких сценариях и в большинстве случаев программная система корректно детектировала, отслеживала и распознавала сущности, записывая всю необходимую статистическую информацию в базу данных. На основании чего можно сделать вывод что разработанный прототип, при незначительных доработках, применим в реальных задачах мониторинга и видеоаналитики.

### **Библиографический список**

1. **Климин, А. А.** Использование искусственного интеллекта в автомобилях / А. А. Климин, С. С. Гилева, Н. Е. Отекина // Мир Инноваций. – 2020. – № 4. – С. 48-51.
2. **Короткова, Е. В.** Машинное обучение и нейронные сети. Важный шаг в развитии человечества / Е. В. Короткова // Современные вопросы естествознания и экономики : Сборник трудов III Международной научно-практической конференции, Прокопьевск, 18 марта 2021 года. – Прокопьевск: Филиал Федерального государственного бюджетного образовательного учреждения высшего профессионального образования "Кузбасский государственный технический университет имени Т. Ф. Горбачева" в г. Прокопьевске, 2021. – С. 240-242.
3. **Загинайло, М. В.** Обучение искусственных нейронных сетей с помощью математического аппарата генетических алгоритмов / М. В. Загинайло // Инновации. Наука. Образование. – 2020. – № 12. – С. 384-389.
4. **Васильева, Е. А.** Сравнение алгоритмов распознавания образов для семантического анализа изображений / Е. А. Васильева // Политехнический молодежный журнал. – 2019. – № 3(32). – С. 8

5. **Zhou, F.** Safety Helmet Detection Based on YOLOv5 / F. Zhou, H. Zhao and Z. Nie // 2021 IEEE International Conference on Power Electronics, Computer Applications (ICPECA), 2021, pp. 6-11, doi: 10.1109/ICPECA51329.2021.9362711.
6. **Htet, K. S.** Event Analysis for Vehicle Classification using Fast RCNN /K. S. Htet and M. M. Sein // 2020 IEEE 9th Global Conference on Consumer Electronics (GCCE), 2020, pp. 403-404, doi: 10.1109/GCCE50665.2020.9291978.
7. **Jung, H.** ResNet-Based Vehicle Classification and Localization in Traffic Surveillance Systems / H. Jung, M. Choi, J. Jung, J. Lee, S. Kwon and W. Y. Jung // 2017 IEEE Conference on Computer Vision and Pattern Recognition Workshops (CVPRW), 2017, pp. 934-940, doi: 10.1109/CVPRW.2017.129.
8. **Mehta, S.** CNN based Traffic Sign Classification using Adam Optimizer / S. Mehta, C. Paunwala and B. Vaidya// 2019 International Conference on Intelligent Computing and Control Systems (ICCS), 2019, pp. 1293-1298, doi: 10.1109/ICCS45141.2019.9065537.
9. **Esteves, R. M.** Competitive K-Means, a New Accurate and Distributed K-Means Algorithm for Large Datasets/ R. M. Esteves, T. Hacker and C. Rong// 2013 IEEE 5th International Conference on Cloud Computing Technology and Science, 2013, pp. 17-24, doi: 10.1109/CloudCom.2013.89.
10. **Hargrave, P. J.** A tutorial introduction to Kalman filtering / P. J. Hargrave // IEE Colloquium on Kalman Filters: Introduction, Applications and Future Developments, 1989, pp. 1/1-1/6.

## ЭМПИРИЧЕСКИЙ КРИТЕРИЙ КАЧЕСТВА КЛАССИФИКАТОРОВ В МНОГОМЕРНОМ ПРОСТРАНСТВЕ ПРИЗНАКОВ

А.О. Гурина<sup>1,2</sup>

ORCID: 0000-0003-1209-7324 e-mail: asya.gurina001512@yandex.ru

В.Л. Елисеев<sup>1,2</sup>

ORCID: 0000-0002-9341-7475 e-mail: vlad-eliseev@mail.ru

<sup>1</sup>Национальный исследовательский университет «Московский энергетический институт»<sup>2</sup>АО «ИнфоТеКС», г. Москва

Рассматривается проблема оценки качества классификаторов в многомерном пространстве признаков, имеющем число измерений более трёх. Обучение классификатора на данных высокой размерности увеличивает риск, как переобучения, так и недообучения, что приводит к непредсказуемым результатам классификации. Результат обучения классификаторов в таком пространстве не поддается интерпретации и визуализации, поскольку проверка соответствия областей классов из обучающего набора и областей, ассоциированных с классами после обучения, затруднена. Отсутствие такой проверки не позволяет оценить уязвимость классификатора к ложным срабатываниям в случае переобучения и к состязательным атакам в случае недообучения. Для решения этой проблемы ранее был введен новый критерий качества, включающий четыре метрики (EDCA). В работе продемонстрировано применение критерия качества классификаторов EDCA на общедоступном наборе данных высокой размерности.

**Ключевые слова:** показатели качества, точность обучения, классификация, искусственные нейронные сети, машинное обучение, многомерное пространство признаков, состязательные атаки

## EMPIRICAL CRITERION OF CLASSIFIER QUALITY IN A MULTIDIMENSIONAL FEATURE SPACE

A.O. Gurina<sup>1,2</sup>, V.L. Eliseev<sup>1,2</sup><sup>1</sup>National research university «Moscow power engineering institute»<sup>2</sup>JSC InfoTeCS, Moscow

**Purpose:** This article notes the problems of classifier quality assessment in a multidimensional space. The quality criterion for constructing classifiers resistant to false positives and adversarial attacks in a high-dimensional space is proposed.

**Design/methodology/approach:** The proposed criterion is based on the discrete estimation of data volumes in the feature space. The criterion includes four quality characteristics: Approx, Excess, Deficit, Coating. Characteristics are calculated for each class. For this, we estimate the volume of class from the training set and the volume of data associated with this class in feature space after training the classifier. Thus, the criterion allows us to evaluate the correspondence of the training set and the data set recognized by the trained classifier in the feature space.

**Findings:** The new criterion of classifiers quality is proposed and tested on the example of the public dataset of high dimension.

**Research limitations/implications:** The algorithm for calculating the characteristics of Approx, Excess, Deficit, Coating in the case of high dimensionality is associated with high requirements for RAM. Memory requirements increase with the number of functions. Further research will be aimed at eliminating dimensional constraints.

**Originality/value:** The quality criterion of classifiers is new. The criterion can be applied to check the correspondence of the partition of the feature space by the training set and the partition of the feature space by the trained classifier. Methods for obtaining such an estimate in a high-dimensional space are unknown. Based on the obtained estimates and criteria, it is possible to draw conclusions about the quality of the classifier, as well as to select parameters for constructing the best classifier. The EDCA criterion is designed to synthesize classifiers resistant to false positives and adversarial attacks in a high-dimensional space. The reported study was funded by RFBR according to the research project № 20-37-90073.

**Keywords:** quality characteristics, learning accuracy, classification, artificial neural networks, machine learning, multidimensional feature space, adversarial attacks



## Введение

Благодаря широкому распространению методов машинного обучения стали известны многочисленные случаи парадоксальных результатов классификации в окрестности точек обучающего множества [1] или зашумленных данных целевого класса [2], на которых, казалось бы, ошибок быть не должно. Такая особенность классификаторов не только негативно отражается на качестве классификации, но и является серьезной уязвимостью, которую используют злоумышленники для целенаправленных атак на системы искусственного интеллекта.

Атаки, основанные на формировании специального примера входных данных, который вызывает ошибку классификатора, называются состязательными, а такие примеры – состязательными примерами. Обычно состязательные примеры определяются через минимальное отклонение от корректно классифицируемых примеров [2]. Так, для атак на системы распознавания и классификации изображений злоумышленники воздействуют на входные данные путем минимального искажения изображения, например, изменения одного пикселя или добавления шума [2]. Широко известны примеры реальных атак и на другие системы: атака на голосового ассистента [4], обход анти-спам фильтров [5], атаки на системы компьютерного зрения [6]- [7] и системы распознавания лиц [8] и другие [9]. Известные прецеденты ошибок методов машинного обучения не позволяют полностью полагаться на этот инструмент при решении задач с высокой ценой ошибки. К таким системам можно отнести автопилот. Ошибки встроенного в автопилот автомобиля марки Tesla классификатора уже стали резонансными [10], [11], [12] как и последствия этих ошибок на дороге [13]. Все эти примеры подтверждают уязвимость интеллектуальных классификаторов любых входных данных.

Поскольку большинство перечисленных уязвимых классификаторов внедрены в эксплуатирующиеся системы и сервисы, нет сомнений в том, что они прошли тщательное тестирование и были высоко оценены по существующим критериям качества. Что позволяет сделать вывод, что оценка классификатора по традиционным критериям качества никак не отражает реальную уязвимость классификатора к ошибкам и состязательным атакам. Рассмотрим традиционные критерии качества классификаторов более подробно.

## Анализ

Контроль качества интеллектуальных классификаторов осуществляется с помощью критериев, основанных на анализе матрицы ошибок (confusion matrix) и ряда метрик (*Accuracy*, *Precision*, *Recall*, *F – score*), рассчитываемых на её основе [14]. Методология оценки качества бинарного классификатора включает анализ матрицы ошибок вида, показанного в таблице 1.

Таблица 1.

Матрица ошибок

Класс \ Классификация	Positive	Negative
Positive	<i>TP</i>	<i>FP</i>
Negative	<i>FN</i>	<i>TN</i>

В ячейках таблицы указывается количество фактов корректной (True) и некорректной (False) классификации для положительного (Positive) и отрицательного (Negative) класса тестовой выборки.

Традиционно, выводы о качестве классификатора делаются на основе характеристик, рассчитываемых по 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:

$$F_1 = 2 \frac{Precision \cdot Recall}{Precision + Recall}$$



Для случая трех и более классов характеристики, изначально разработанные для оценки бинарной классификации, усредняются для оценки многоклассового классификатора. Также, из формул следует, что значения характеристик зависят от состава и сбалансированности тестовой выборки. Ранее отмечалось, что проблема несбалансированности может быть устранена при условии задания специфической структуры классификатора [15]. В обзоре методов работы с несбалансированными данными [16] предлагают использовать такие показатели, как коэффициент каппа Коэна ( $\kappa$ ) или коэффициент корреляции Мэтьюза (MCC), поскольку они менее чувствительны к дисбалансу размеров классов.

Все отмеченные показатели основаны на классах, описываемых ограниченными счётными множествами, что хорошо согласуется с практическими реализациями, оперирующими тестовыми и контрольными эмпирическими выборками. Однако следует отметить, что именно это и не позволяет им дать объективную оценку свойств и качества классификатора для работы с данными за пределами выборки, используемой для обучения и тестирования. При этом классификатор, получая высокую оценку качества на основе оценки качества классификации ограниченной тестовой выборки, применяется для любых входных данных, в том числе находящихся за пределами области обучающего и тестового множества. В тоже время, совершенно очевидно, что после корректного обучения классификатор способен распознавать примеры, которых не было в обучающей или тестовой выборке. Это ценное свойство алгоритмов машинного обучения приводит к тому, что в пространстве признаков область, распознаваемая обученным классификатором, может сильно отличаться от истинной области класса. Однако никакого контроля за степенью и корректностью обобщения классификатором знаний, полученных из ограниченной обучающей выборки, текущие подходы по оценке качества не предусматривают. Этот факт существенно снижает доверие к результатам оценки и объясняет вышеописанные случаи непредвиденной классификации, а также делает менее предсказуемым и надежным применение классификатора, оцененного таким образом, в прикладных задачах.

Востребованность алгоритмов машинного обучения и распространенность состязательных атак делает традиционный подход уже недостаточным для оценки качества классификаторов в виду отсутствия контроля свойств классификатора за пределами известных примеров.

Анализ литературы позволяет сделать вывод о том, что непредсказуемое поведение интеллектуальных классификаторов и недостаточность существующих критериев делает вопрос разработки критерия качества классификаторов, учитывающего свойства классификатора за пределами обучающей и тестовой выборки, весьма актуальным.

### Методология

Ранее для оценки качества одноклассовых классификаторов был введен и успешно апробирован многомерный критерий EDCA [17], позволяющий численно охарактеризовать свойства интеллектуального классификатора за пределами известных выборок. Основным отличием критерия EDCA от традиционных является независимость оценки от размера и состава тестовой выборки, поскольку критерий основан на анализе поведения обученного классификатора в окрестностях обучающего множества в пространстве признаков. Для случая одного класса, качество классификатора описывалось четырьмя характеристиками: *Excess*, *Deficit*, *Coating*, *Approx*. Совокупность этих характеристик позволяет оценить соответствие области обучающих данных и области фактически распознаваемых обученным классификатором данных в пространстве признаков.

Для расчета характеристик необходимы две дискретные оценки: оценка объема обучающих данных  $|X_T|$  в пространстве признаков и оценка объема фактически распознаваемых обученным классификатором данных  $|X_D|$  в пространстве признаков. Рассмотрим способ их определения. Для этого введем в рабочей области векторного пространства  $R^n$ , 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  $a$ :

$$\forall x \in R^n \exists ! a | x \in a$$

На практике нам доступно конечное счётное обучающее множество  $X_T$ . Тогда объем этого множества будет равен сумме объемов элементов разбиения, включающих хотя бы одну точку

множества  $X_T$ , или количеству таких элементов, умноженному на единичный объем элемента разбиения, в случае разбиения пространства на ячейки одинакового объема.

Таким образом, дискретное множество  $X_T^*$  образуется из элементов разбиения, содержащих хотя бы одну точку из  $X_T$ :

$$X_T^* = \{a | \exists x \in X_T \& x \in a\}$$

Обозначим отнесение точки  $x$  обученным классификатором к целевому классу «1» как  $CL^1(x) = 1$ . Тогда, дискретное множество  $X_D^*$  образуется из элементов разбиения, какая-либо точка внутри которых дает положительный результат классификации:

$$X_D^* = \{a | \exists x \in a: CL^1(x) = 1\}$$

Дискретная оценка объема обучающего множества  $|X_T|$  рассчитывается через сумму объемов атомарных элементов разбиения, образующих множество  $X_T^*$ . Дискретная оценка объема обучающего множества  $|X_D|$  рассчитывается через сумму объемов атомарных элементов разбиения, образующих множество  $X_D^*$ . 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  $X_D^*$ .

После определения описанных объёмов, рассчитать характеристики *Excess*, *Deficit*, *Coating*, нормированные на объём целевого множества, и характеристику *Approx*, можно по следующим формулам:

$$Excess = \frac{|X_D \setminus X_T|}{|X_T|} \quad Deficit = \frac{|X_T \setminus X_D|}{|X_T|} \quad Coating = \frac{|X_T \cap X_D|}{|X_T|} \quad Approx = \frac{|X_T|}{|X_D|}$$

Характеристика *Excess* (англ. избыток) принимает значение больше 0 в том случае, если  $CL^1$  неправильно классифицирует объекты за пределами целевого множества. Чем больше область распознавания выходит за пределы обучающего множества, тем больше будет значение *Excess*. Характеристика *Deficit* (англ. недостаток) принимает значение от 0 до 1, причем, если  $CL^1$  неправильно классифицирует объекты в пределах целевого множества, то характеристика больше 0. Предельным значением  $Deficit = 1$  является ситуация, когда целевое и распознаваемое множества не пересекаются. Характеристика *Coating* принимает значения от 0 до 1, причем значению 0 соответствует нулевое перекрытие множеств, а 1 соответствует полное покрытие целевого множества. Характеристика *Approx* показывает точность аппроксимации целевого множества и принимает значение 1, когда объем распознаваемой обученным классификатором области равен объему области обучающего множества в пространстве признаков, и стремится к 0, когда фактически распознаваемый объем существенно больше объема обучающих данных.

Полное соответствие распознаваемой обученным классификатором области и области обучающего множества в пространстве признаков характеризуется следующими значениями характеристик:

$$Excess = 0, \quad Deficit = 0, \quad Coating = 1, \quad Approx = 1 \quad (1)$$

Рассмотрим критерий оценки EDCA, обобщенный для случая, когда число классов больше одного, на примере оценки классификатора, обученного на обучающем наборе данных  $X_T$ , включающем  $M$  классов. Обозначим множество точек из обучающего набора, соответствующих классу  $i$  как  $X_T^{(i)}$ , где  $i = 1, \dots, M$ . Множество точек, которое обученный классификатор распознает как класс  $i$ , обозначим как  $X_D^{(i)}$ .

Для расчета предложенных показателей качества необходимо ввести сеточную аппроксимацию непрерывного пространства в некоторой рабочей области, в которой будем проводить расчеты характеристик. Ограничив пространство рабочей областью, включающей множества обучающих и тестовых данных, можно нормировать координаты в диапазон от 0 до 1. Введем равномерную сетку с шагом  $h$  по всем координатам ограниченного сеточного пространства. Таким образом, для  $N$ -мерного пространства признаков будем иметь  $\left(\frac{1}{h}\right)^N$  ячеек сеточного пространства. Назовем множество этих ячеек скан-множеством  $X_S$ . Будем относить точки множеств  $X_T$  и  $X_D$  к ячейкам скан-множества, которым они принадлежат. Таким образом,

ячейки обучающих данных образуют множество  $X_T^* \subset X_S$ , а множество  $X_D^* \subset X_S$  будем формировать из ячеек с положительным результатом классификации для точки в центре ячейки.

Размер ячейки рекомендуется подбирать для исследуемого набора данных так, чтобы количество точек обучающего набора в каждой ячейке было сбалансировано и стремилось к 1. Если выбирать размер ячейки слишком малым, то в случае высокой размерности данных расчеты потребуют существенно больше времени. Кроме того, слишком малое значение  $h$  приведет к тому, что соседние точки  $X_T$  окажутся разделены в скан-множестве несколькими пустыми ячейками, не принадлежащими  $X_T^*$ , что может привести к потере значимости численных оценок EDCA.

Для численной оценки свойств многоклассового классификатора может быть использован многомерный критерий EDCA, включающий введенные ранее характеристики для каждого из классов. В этом случае, для каждого из классов могут быть рассчитаны свои характеристики  $Excess^{(i)}$ ,  $Deficit^{(i)}$ ,  $Coating^{(i)}$ ,  $Approx^{(i)}$ .

Применение критерия качества, размер которого зависит от числа классов, представляется неудобным. Однако, в отсутствии предпочтений по значимости того или иного класса разумной числовой оценкой качества классификатора можно считать кортеж из 4 значений:  $Excess$ ,  $Deficit$ ,  $Coating$ ,  $Approx$ , каждое из которых наихудшим образом характеризует классификатор по всем классам:

$$Excess = \sum_i Excess^{(i)} \qquad Coating = \min_i Coating^{(i)}$$

$$Deficit = \max_i Deficit^{(i)} \qquad Approx = \min_i Approx^{(i)}$$

Для характеристик  $Deficit$ ,  $Coating$ ,  $Approx$  берется наихудшее значение (минимальное или максимальное), а для  $Excess$  – сумма по всем классам. Совокупность предложенных характеристик образует многомерный критерий качества EDCA, который можно использовать для сравнения классификаторов.

Стоит отметить, что классификатор, характеризующийся предельными значениями характеристик (1), может оказаться уязвим для ошибок классификации второго рода (FN) в случае нерепрезентативной обучающей выборки. Подобные ошибки классификации в некотором роде тоже можно считать constitutive атаками. Однако, если в силу специфики решаемой задачи ошибки первого рода более критичны, чем ошибки второго рода, то рекомендуется выбирать классификатор, характеризующийся значениями, близкими к предельным (1).

Также, для случая бинарной и многоклассовой классификации, важно учитывать, что в пространстве признаков возможно частичное перекрытие областей, содержащих примеры, относящиеся к разным классам. Например, такой особенностью обладает набор данных «Ирисы Фишера»[18]. В таком случае построение идеального классификатора представляется невозможным в принципе, что хорошо известно исследователям машинного обучения.

Характеристикой неоднозначности набора обучающих данных является ненулевое пересечение областей классов  $i$  и  $j$ :  $X_T^{(i)} \cap X_T^{(j)} \neq \emptyset$ . В области неоднозначного определения классификатора любой пример будет в каком-то смысле constitutive. При построении сеточного пространства следует подбирать такой шаг  $h$ , чтобы пересечение дискретных множеств разных классов было минимальным. Кроме того, в случае ненулевого пересечения рекомендуется оценивать объем общих границ областей разных классов, распознаваемых обученным классификатором. При одинаковых характеристиках критерия EDCA классификатор с меньшим объемом общих границ является более устойчивым к ошибкам классификации вблизи общих границ классов.

Предложенный метод оценки классификаторов особенно ценен для многомерного пространства признаков, когда визуализация областей распознавания обученного классификатора затруднена или невозможна.

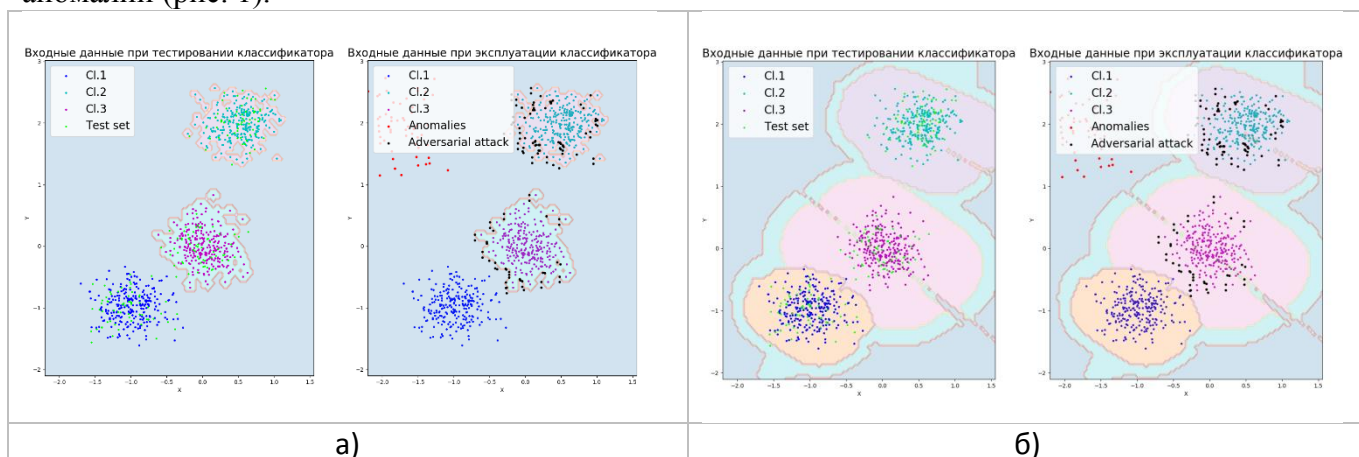
### Эксперименты

Для наглядной демонстрации эффективности предлагаемого критерия был проведен ряд экспериментов, включающих обучение, тестирование и сравнение качества двух классификаторов по традиционным критериям качества и по критерию EDCA. При этом тестирование проводилось

как на тестовой выборке, так и на выборке, включающей состязательные примеры и аномалии для сравнения устойчивости классификаторов и объективности исследуемых критериев качества.

В данной работе состязательные примеры представляли собой примеры обучающей выборки с минимальным отклонением в значении одного из признаков, вызывающие ошибки классификатора. В экспериментах первый классификатор был реализован традиционными методами машинного обучения, такими как SVM или нейронная сеть DNN типа многослойный персептрон, а второй классификатор – на основе ансамбля автокодировщиков (EnAE), который впервые был предложен для динамической классификации в условиях дрейфа концепций [19]. Ансамбль автокодировщиков в качестве второго классификатора выбран не случайно, а исходя из предположения, что он более устойчив к ошибкам в виду особенности формируемых после обучения областей распознавания – область распознавания каждого класса замкнута, а компактность области легко регулируется с помощью порогов ошибки реконструкции (рис. 1б). Такое свойство обеспечивается потому, что один автокодировщик ансамбля обучается распознавать данные всех известных классов, а каждый из остальных автокодировщиков ансамбля обучается распознавать данные только одного из классов.

Рассмотрим результаты обучения SVM-классификатора и ансамбля автокодировщиков EnAE на синтетическом двумерном наборе, включающем три класса, а также результаты тестирования, как на тестовом наборе, так и на наборе, включающем состязательные примеры и аномалии (рис. 1).



**Рис. 1. Фактические области распознавания и результаты тестирования классификаторов SVM(а) и EnAE(б) для синтетического двумерного набора**

Видно, что как на рис. 1а, так и на рис. 1б примеры тестовой выборки классифицируются SVM и EnAE верно, однако большинство состязательных атак и аномалий обученный SVM ошибочно относит к «1» классу (рис. 1а справа), когда как обученный EnAE классифицирует эти примеры без ошибок (рис. 1б справа).

Рассмотрим оценки качества классификаторов по традиционным характеристикам (*Precision*, *Recall*, *F1*) на тестовой выборке и по критерию EDCA, приведенные в табл. 2.

**Таблица 2.**

**Оценка качества SVM и EnAE по традиционным критериям и по критерию EDCA**

Классификатор	Традиционные оценки			Оценка по EDCA			
	Precision	Recall	F <sub>1</sub>	Excess	Deficit	Coating	Approx
SVM	0.99	0.98	0.98	37.2	0.00	1.00	0.00
EnAE	0.99	0.98	0.98	5.80	0.00	1.00	0.20

Как видно из табл. 2 качество классификаторов по традиционным показателям, рассчитанным для тестовой выборки, одинаково, в то время как значения характеристик *Excess* и *Approx* для SVM и EnAE сильно отличаются. Согласно предлагаемой методологии, оценки классификатора EnAE ближе к предельным значениям (1), характеризующим полное соответствие объема обучающих данных и объема фактически распознаваемых данных в пространстве

признаков, а сам классификатор более надежен, чем исследуемый SVM. Количество ошибочно классифицированных исследуемыми классификаторами состязательных примеров и аномалий приведено в табл. 3.

Таблица 3.

Количество ошибок классификации SVM и EnAE на выборке состязательных примеров и аномалий

Классификатор	Состязательные атаки	Аномалии
SVM	120	70
EnAE	1	0

Полученные результаты тестирования исследуемых классификаторов в контексте уязвимости к состязательным примерам и аномалиям согласуются с оценкой, полученной по критерию EDCA.

Рассмотрим результаты апробации критерия EDCA на примере оценки качества классификаторов общедоступных многомерных наборов данных. Сравним качество классификаторов SVM и EnAE на четырехмерном наборе данных «Ирисы Фишера» [18]. Разделим исходный набор данных на обучающее (80%) и тестовое (20%) множество. Обучающий набор данных представлен в сжатом пространстве с помощью алгоритма визуализации многомерных данных - T-SNE (рис. 2).

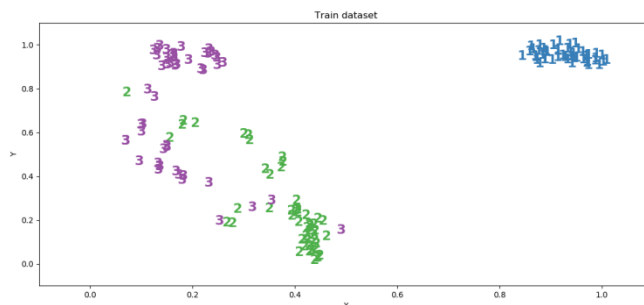


Рис. 2. Обучающий набор данных «Ирисы Фишера»

Синтез классификатора SVM проводился с параметрами  $\gamma = 100$ ,  $C = 10$ . В классификаторе EnAE все четыре автокодировщика имели типовую архитектуру  $NN_{4,4,3,2,3,4,4}$  и обучались на протяжении  $3 \cdot 10^5$  эпох с помощью алгоритма обучения ADAM.

Рассчитанные традиционные характеристики качества на тестовой выборке и характеристики EDCA для обоих классификаторов представлены в таблице 4. Для расчетов скан-множество было определено с шагом  $h = 0.14$ . Также для проверки фактической устойчивости классификаторов был сгенерирован второй тестовый набор, включающий 3280 размеченных примеров с минимальным отличием от примеров обучающей выборки. В результате классификации этих примеров, было подсчитано количество успешных состязательных атак на каждый классификатор (табл. 4).

Таблица 4.

Оценка качества SVM и EnAE по традиционным критериям и по критерию EDCA для набора «Ирисы Фишера»

Классификатор	Традиционные оценки			Оценка по EDCA				Состязательные атаки
	Precision	Recall	F <sub>1</sub>	Excess	Deficit	Coating	Approx	
SVM	0.97	0.97	0.97	93.94	0.00	1.00	0.01	115
EnAE	1.00	0.97	0.98	16.82	0.00	1.00	0.08	18

Анализ значений традиционных оценок, показывает, что качество классификаторов не сильно отличается, а значения характеристик Excess показывают, что фактическая область распознавания SVM-классификатора существенно превосходит область распознавания EnAE и область обучающих данных. Согласно предлагаемому критерию, качество данного EnAE-классификатора превосходит данный SVM-классификатор, что подтверждается меньшим количеством ошибок классификации состязательных примеров.

Примеры обучающей выборки и успешные состязательные атаки на классификатор SVM показаны в сжатом пространстве признаков на Рис., а на классификатор EnAE на рис.4.

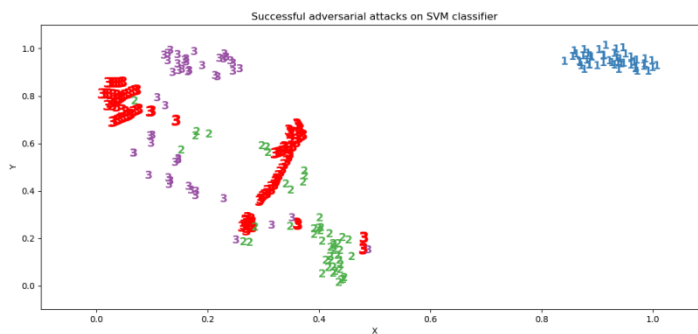


Рис.3. Успешные состязательные атаки на классификатор SVM (выделены красным)

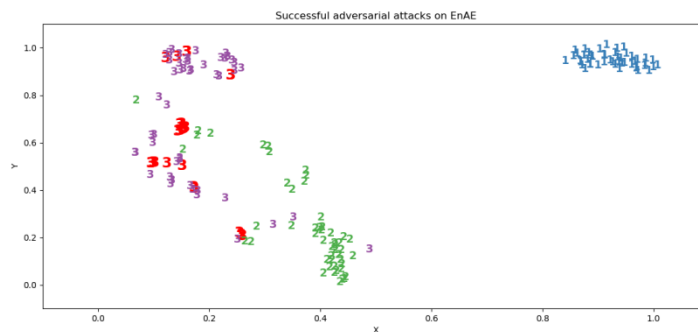


Рис.4. Успешные состязательные атаки на классификатор EnAE (выделены красным)

Полученные результаты классификации второго тестового набора подтверждают, что предлагаемый критерий EDCA более информативен в контексте оценки качества классификаторов за пределами обучающей и тестовой выборки, чем традиционные характеристики.

Возьмем другой набор данных «Семена пшеницы» [20] и повторим эксперименты по тестированию и сравнению двух видов классификаторов: глубокая нейронная сеть DNN и ансамбль автокодировщиков EnAE. В этом наборе присутствуют данные трех классов, заданные векторами признаков с размерностью 7. На рис. 5 представлена двумерная визуализация обучающего множества с помощью алгоритма T-SNE.

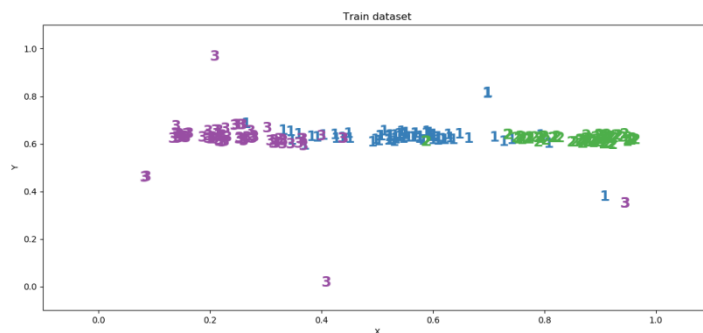


Рис.5 Обучающий набор данных «Семена пшеницы»

Нейронная сеть DNN включала 4 полносвязных слоя, слой Dropout и выходной слой Softmax из 3-х нейронов. Архитектура сети DNN представлена ниже:

$$y = NN_{7,7,D,5,D,3,D,3}(x), \quad x \in R^7, y \in R^3$$

Обучение DNN проводилось в течение  $5 * 10^4$  эпох.

Также для этого набора данных был синтезирован классификатор EnAE, включающий 4 нейросетевых автокодировщика. Архитектура каждого автокодировщика указана ниже:

$$y = NN_{7,7,6,4,3,4,6,7}(x), \quad x, y \in R^7$$

Сравним качество нейросетевого классификатора DNN и классификатора EnAE по традиционным критериям качества и по критерию EDCA, проанализировав значения, представленные в табл. 5, и соотнесем их с количеством успешных состязательных атак на классификаторы. В алгоритме расчета характеристик EDCA шаг сетки разбиения пространства признаков выбран равным 0.08. Второй тестовый набор включал 1960 размеченных примеров, минимально отличающихся от примеров обучающей выборки. После классификации этих примеров, было подсчитано количество успешных состязательных атак на каждый классификатор (табл. 5).

По традиционным критериям качество обоих классификаторов достаточно высокое, при этом NN незначительно превосходит EnAE. Сравнение значений характеристик Excess для исследуемых классификаторов показывает, что у классификатора DNN область распознавания за пределами обучающей области существенно выше, чем у EnAE. Согласно предлагаемому критерию это позволяет предположить, что классификатор DNN подвержен ошибкам больше, чем EnAE. Данное предположение подтверждается результатами эксперимента, приведенными в табл. 5: на классификатор DNN было успешно проведено более чем в 17 раз больше состязательных атак, чем на классификатор EnAE.

Таблица 5.

Оценка качества DNN и EnAE по традиционным критериям и по критерию EDCA для набора «Семена Пшеницы»

Классификатор	Традиционные оценки			Оценка по EDCA				Состязательные атаки
	Precision	Recall	F <sub>1</sub>	Excess	Deficit	Coating	Approx	
SVM	1.00	1.00	1.00	735998.84	0.10	0.90	0.00	759
EnAE	1.00	0.98	0.99	410.80	0.70	0.30	0.06	43

Для наглядности на *Рис.1* показана обучающая выборка и успешные состязательные атаки на классификатор DNN, а на *Рис.7* – на классификатор EnAE. Красным выделены метки классов, распознанные не так, как размечены соседние с ними примеры обучающего набора данных.

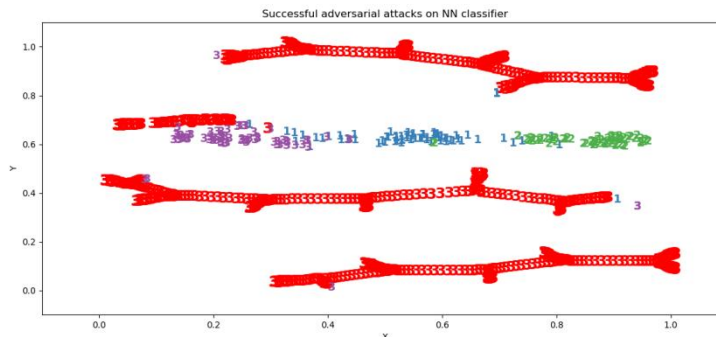


Рис.1. Успешные состязательные атаки на классификатор DNN (выделены красным)

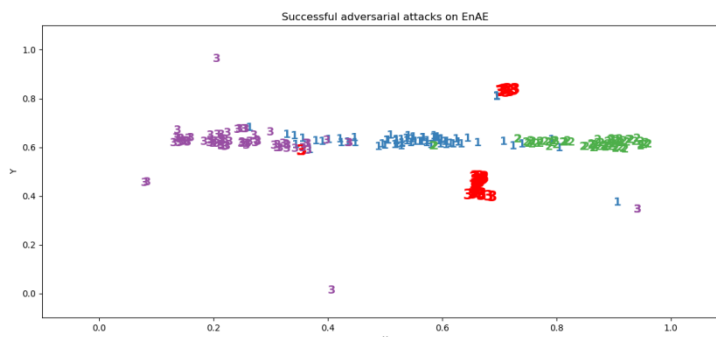


Рис.7. Успешные состязательные атаки на классификатор EnAE (выделены красным)

Сопоставляя результаты тестирования и оценки качества исследуемых классификаторов по традиционным критериям и по критерию EDCA можно убедиться, что многомерный критерий



EDCA позволяет подробнее охарактеризовать качество классификатора и в области обучающей выборки, и за её пределами.

### Обсуждение

Предлагаемый многомерный критерий EDCA позволяет сравнивать классификаторы по точности соответствия их фактической области распознавания и области обучающих данных в пространстве признаков.

Отметим преимущества предлагаемого многомерного критерия качества EDCA:

- работает для любой реализации классификатора;
- не требуется тестовая выборка;
- не чувствителен к дисбалансу классов;
- позволяет численно оценить устойчивость фактической области распознавания классификатора к состязательным атакам и ошибкам в многомерном пространстве, когда визуализация затруднена.

Многомерный критерий EDCA имеет следующие недостатки:

- чувствительность к выбору размера ячейки разбиения пространства признаков;
- высокие требования к оперативной памяти для расчетов в многомерном пространстве;
- полиномиальная сложность вычислений со степенью полинома, равной размерности пространства признаков, что делает невозможным прямое применение критерия для задач с большой размерностью, в частности, при анализе изображений.

Следует отметить, что, несмотря на перечисленные недостатки, альтернативных методов для оценки свойств поведения классификаторов за пределами обучающей выборки не было найдено. В то же время, представляется возможным оптимизировать процедуру расчета EDCA как с помощью алгоритмических подходов, так и аппаратных средств (GPGPU). Дальнейшие исследования могут привести к методике по расчету оптимального шага для конкретного набора данных.

Включение предлагаемого критерия в методологию оценки качества классификаторов должно повысить доверие к получаемым оценкам. Применение предлагаемого критерия для оценки классификаторов, также позволит настраивать классификаторы до более устойчивых состояний, что повысит доверие к методам машинного обучения и расширит область их применения.

### Заключение

В данной работе рассматривается проблема непредсказуемости методов машинного обучения и отмечается недостаточность традиционных критериев для оценки качества классификатора за пределами известных выборок. Предлагаемый многомерный критерий качества классификаторов EDCA основан на принципиально новом подходе к оценке и позволяет в пространстве признаков оценивать соответствие фактических областей распознавания классификатора обучающим данным. Критерий был апробирован для оценки многоклассовых классификаторов двумерных и многомерных данных. Результаты экспериментов подтверждают, что многомерный критерий EDCA позволяет охарактеризовать качество классификатора как в области обучающей выборки, подобно традиционным критериям, так и за её пределами, что недоступно для традиционных критериев качества.

### Примечание

Работа проведена при финансовой поддержке РФФИ, проект № 20-37-90073.

### Библиографический список

1. **Szegedy, C., Liu, W., Jia, Y., Sermanet, P., Reed, S., Anguelov, D., Erhan, D., Vanhoucke, V., Rabinovich, A.** Going deeper with convolutions. //In 2015 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Boston, MA, 2015.
2. **Goodfellow, J. Shlens, C.** Szegedy.Explaining and harnessing adversarial examples //arXiv 1412.6572., 2014.
3. **Su, J., Vargas, D., Sakurai, K.** One Pixel Attack for Fooling Deep Neural Networks // IEEE Transactions on Evolutionary Computation, vol. 23, no. 5, pp. 828-841, 2019.



4. **Agrafiotis, M., Goldsmith, I., Bispham, M.** Nonsense Attacks on Google Assistant and Missense Attacks on Amazon Alexa. // In Proceedings of the 5th International Conference on Information Systems Security and Privacy - ICISSP, 2019.
5. **Ozkan, H, Sen, S., Burcu, C.** Analysis of Adversarial Attacks against Traditional Spam Filters // In Proceedings of International Conference on All Aspects of Cyber Security, 2019.
6. **Harris, M.** Researchers find a malicious way to meddle with autonomous cars. University of Washington, 2017. [Электронный ресурс]. URL: <https://www.caranddriver.com/news/a15340148/researchers>.
7. **Kurakin, A. Goodfellow, I., Bengio, S.** Adversarial examples in the physical world. // ICLR 2017, 2017.
8. **Sharif, SMahmood, Bhagavatula, Sruti, Bauer, Lujo, Reite, Michael K.** Accessorize to a Crime: Real and Stealthy Attacks on State-of-the-Art Face Recognition // In Proceedings of the 2016 ACM SIGSAC Conference on Computer and Communications Security (CCS '16). Association for Computing Machinery, New York, NY, USA, 1528–1540.
9. **Wiggers, K.** Adversarial attacks in machine learning: What they are and how to stop them.2021. [Электронный ресурс]. URL: <https://venturebeat.com/2021/05/29/adversarial-attacks-in-machine-learning-what-they-are-and-how-to-stop-them/>.
10. **Povolny, S.** Model Hacking ADAS to Pave Safer Roads for Autonomous Vehicles.2020. [Электронный ресурс]. URL: <https://www.mcafee.com/blogs/other-blogs/mcafee-labs/model-hacking-adas-to-pave-safer-roads-for-autonomous-vehicles/>.
11. **Ackerman, E.** Three Small Stickers in Intersection Can Cause Tesla Autopilot to Swerve Into Wrong Lane. 2019. [Электронный ресурс]. URL: <https://spectrum.ieee.org/three-small-stickers-on-road-can-steer-tesla-autopilot-into-oncoming-lane>.
12. Tencent Keen Security Lab. Adversarial Machine Learning against Tesla's Autopilot.2019. [Электр. ресурс]. URL: [https://www.schneier.com/blog/archives/2019/04/adversarial\\_mac.html](https://www.schneier.com/blog/archives/2019/04/adversarial_mac.html).
13. **Lambert, F.** Understanding the fatal Tesla accident on Autopilot and the NHTSA probe.2016. [Электронный ресурс]. URL: <https://electrek.co/2016/07/01/understanding-fatal-tesla-accident-autopilot-nhtsa-probe/>.
14. **Forman, G.** An extensive empirical study of feature selection metrics for text classification // Journal of Machine Learning Research (JMLR), no. 3, 2003.
15. **Lee, W., Liu, B.** Learning with positive and unlabeled examples using weighted logistic regression // In Proceedings of the Twentieth International Conference on International Conference on Machine Learning (ICML'03), 2003.
16. **Haixiang, G., Yijing, L., Shang, J., Mingyun, G., Yuanyue, H., Bing, G.** Learning from class-imbalanced data: Review of methods and applications// Expert Systems with Applications, vol. 73, pp. 220-239, 2017.
17. **Гурина, А., Елисеев, В.** Эмпирический критерий качества одноклассового классификатора //Сборник материалов XXVII Международной научно-технической конференции Нижегородский государственный технический университет им. Р.Е. Алексева. Нижегородский государственный технический университет им. Р.Е. Алексева. Нижний Новгород, 2021. С. 648-657.
18. **R. Fisher.** Irisdataset. [Электронный ресурс]. URL: <http://archive.ics.uci.edu/ml/datasets/Iris>.
19. **Gurina, A., Eliseev, V., Kolpinskiy, S.** Dynamic classification approach using scalable ensemble of autoencoders to classify data with drift // Journal of Physics: Conference Series, vol. 2134, no. 1, 2021.
20. Seeds dataset [Электронный ресурс]. URL: <http://archive.ics.uci.edu/ml/datasets/seeds>.

Задача классификации является одной из классических задач, для решения которых на практике применяют аппарат искусственных нейронных сетей (ИНС). В последние годы исследователями было предложено и реализовано множество разнообразных архитектур нейросетевых классификаторов, в том числе основанных на глубоких нейронных сетях. Однако вопросы обоснованного выбора той или иной архитектуры и формальной оценки сложности нейронной сети, требуемой для решения задачи, в каждом конкретном случае остаются открытыми. Обоснованный выбор архитектуры и оценка необходимой сложности нейронной сети в конкретном случае позволят экономить время и вычислительные мощности при решении задач классификации. В статье проводится аналитический обзор архитектур искусственных нейронных сетей, используемых для решения задач классификации. Анализируется степень формализации подходов, используемых для синтеза архитектур глубоких ИНС. Рассматриваются основные области применения нейросетевых классификаторов. Делаются выводы о виде наиболее подходящих задач, для которых возможна разработка обоснованных методов синтеза архитектуры ИНС.

**Ключевые слова:** архитектура искусственных нейронных сетей, задача классификации, многослойный персептрон, сверточные нейронные сети, глубокие нейронные сети.

## OVERVIEW OF METHODS FOR THE SYNTHESIS OF NEURAL NETWORK CLASSIFIERS ARCHITECTURE

S.V. Kolpinskiy<sup>1,2</sup>, V.L. Eliseev<sup>2,1</sup>

<sup>1</sup>National research university «Moscow power engineering institute»

<sup>2</sup>JSC InfoTeCS, Moscow

**Purpose:** This article provides an analytical review of artificial neural network architectures used for the solution of various classification problems.

**Design/methodology/approach:** The classification problem is one of the classical problems for which artificial neural networks are applied in practice. In recent years, researchers have proposed and implemented many different architectures of neural network classifiers, including highly developed deep neural networks. However, the problems of the reasonable choice of a particular architecture and the formal estimation of the neural network complexity required to solve an exact classification task are still open in each case.

**Findings:** The review consists of a taxonomy and a descriptive classification of the modern neural networks architectures and their subparts.

**Research limitations/implications:** The article provides a review of the most important neural network architectures and their informal description.

**Originality/value:** A reasonable architecture choice and an estimation of the necessary neural network complexity when solving a specific task will save time and computing power when solving classification problems.

**Keywords:** artificial neural network architecture, classification problem, multilayer perceptron, convolutional neural networks, deep neural networks.

## Введение

В настоящее время бурно развиваются и широко применяются на практике методы обработки данных, основанные на машинном обучении, видную роль в которых играет аппарат искусственных нейронных сетей (ИНС). Под понятием искусственной нейросети обычно понимается математическая модель, построенная по аналогии с принципом функционирования биологических нейронных сетей, и призванная «обучаться», то есть настраивать свои параметры путем извлечения опыта из выборочных данных и обобщения этого опыта для последующей обработки новых данных. Аппарат искусственных нейронных сетей зародился в середине XX века, и с тех пор исследователями было приложено большое количество усилий по его развитию и формализации. Были получены теоретические и практические результаты, демонстрирующие эффективность применения нейросетей при решении множества прикладных задач.

К задачам, решаемым с помощью нейросетей, среди прочих, относятся такие задачи как распознавание образов (классификация, обнаружение аномалий, обнаружение и распознавание объектов на изображениях и видео, распознавание звуков и т. п.), прогнозирование, управление, трансформация данных (автоматический перевод текстов, обработка изображений), сжатие данных. Глобально все перечисленные задачи объединяет то, что нейронные сети, за счет своей способности извлекать информацию из данных и обобщать ее, используются для аппроксимации сложных, часто нелинейных, неявных зависимостей, выраженных в выборочных данных.

При построении модели ИНС при решении конкретной прикладной задачи требуется определиться с ее архитектурой. Существует ряд архитектурных идей, на практике доказавших свою применимость [1]. Наиболее плодотворной из них является идея объединения формальных нейронных элементов в так называемые «слои», при этом распространение сигналов в сети происходит от слоя к слою. Такие сети называют слоистыми. К слоистым сетям относятся:

- слоисто-полносвязные сети (в них нейроны соседних слоев соединяются по принципу «каждый с каждым», наиболее известным представителем таких сетей является многослойный персептрон);
- сверточные сети (в них используется операция дискретной свертки данных с матрицами весовых коэффициентов, а нейронные элементы в слоях топологически имеют определенную регулярную структуру);
- сети с обратными связями или рекуррентные сети (в них распространение сигналов может происходить не только от входного слоя к выходному, но и в обратном направлении).

Также существует ряд соображений по применению того или иного вида сетей в зависимости от решаемой задачи. Так, сверточные сети применяются для обработки изображений или любых данных, в которых прослеживается определенная пространственная структура. Рекуррентные сети применяют, когда предполагается, что обрабатываемые данные имеют внутренние зависимости между последовательными элементами (временные ряды, текст, речь и т. п.). Слоисто-полносвязная архитектура многослойного персептрона в этом смысле является базовой и в общем случае применимой к произвольным данным, но уступает по эффективности более узконаправленным видам сетей.

Не смотря на наличие данных архитектурных идей и соображений по их применению, они не носят конкретного характера и не отвечают на вопросы обоснованного выбора количества слоев, количества и типа нейронных элементов в слоях при решении конкретной прикладной задачи. Архитектурные решения чаще всего основываются на интуиции и опыте разработчика нейронной сети или на опыте других разработчиков при решении схожих задач. При этом выбор в каком-то смысле минимальной по количеству нейронных элементов архитектуры позволит сокращать время и экономить вычислительные ресурсы при обучении нейросетей. В некоторых случаях использование избыточных по количеству параметров

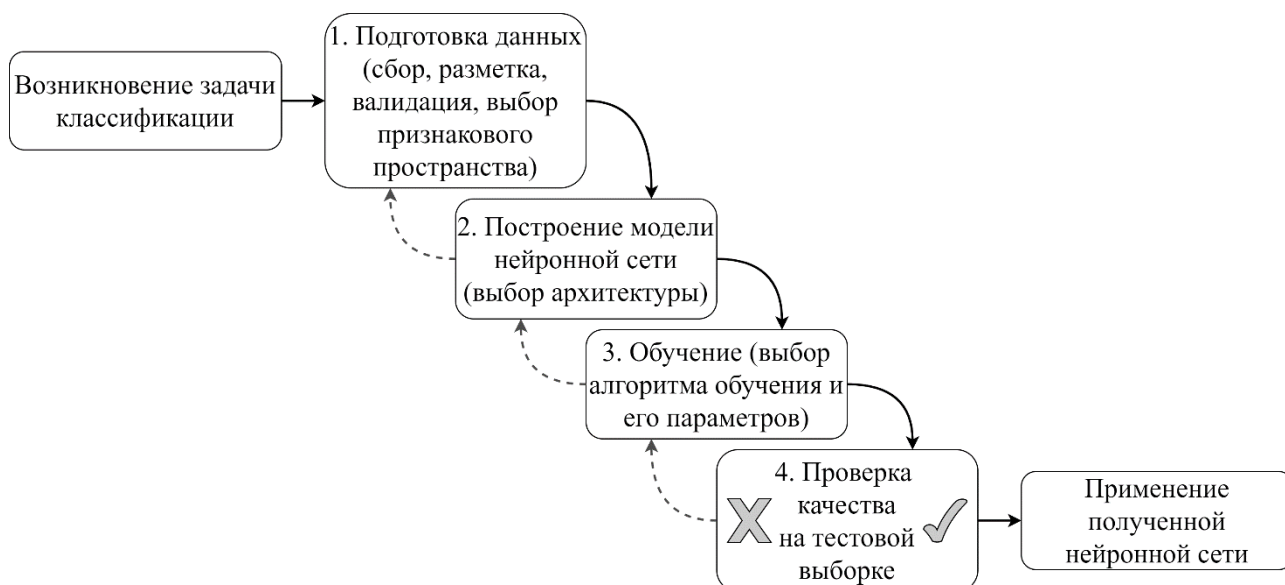
нейросетей оказывается недопустимо (например, при обработке конфиденциальных данных в недоверенной среде, когда сам алгоритм обработки оказывается «растворен» в архитектуре нейросети, наличие избыточного количества параметров в нейросети ставит данные и алгоритм обработки под угрозу раскрытия [2]). С задачей синтеза минимальных архитектур искусственных нейронных сетей тесно связана проблема атак на нейросетевые классификаторы. Основная суть таких атак – в поиске парадоксальных и «неправильных» решений нейросетевого классификатора. Например, для системы машинного зрения беспилотного автомобиля можно сформировать безобидное на первый взгляд изображение, которое будет нейронной сетью восприниматься как дорожный знак, что приведет к неправильному и потенциально опасному решению системы управления автомобилем. Причина таких атак кроется в чрезмерно широком пространстве, относимом к классу визуального объекта. Как правило, это является следствием либо недостаточного, либо чрезмерного количества параметров нейронной сети.

В связи со всем вышесказанным возникает желание сделать шаг в направлении обоснованного выбора архитектуры нейросети для решения конкретной задачи, хотя бы для некоторого класса задач и архитектур. Замечательным результатом в смежной области методов машинного обучения, на который хотелось бы ориентироваться, является понятие размерности Вапника-Червоненкиса [3], показывающее сложность алгоритма бинарной классификации методом SVM для заданного обучающего множества. Целью данной работы является рассмотрение архитектур искусственных нейронных сетей, используемых для решения задач классификации, анализ подходов, используемых для синтеза архитектур ИНС, и попытка сделать предположение о виде наиболее подходящих задач, для которых возможна разработка обоснованных методов синтеза архитектуры ИНС.

### **Нейросетевая классификация**

Классификация – один из разделов машинного обучения, посвященный решению следующей задачи. Имеется множество объектов, разделённых некоторым образом на классы. Множество классов конечно и определено заранее. Задано конечное множество объектов, для которых известно, к каким классам они относятся. Это множество называется обучающей выборкой. Классовая принадлежность остальных объектов неизвестна. Требуется построить алгоритм (классификатор), способный классифицировать произвольный объект из исходного множества. Классифицировать объект – значит указать номер (или наименование) класса, к которому классификатор относит данный объект. Классификация объекта – номер или наименование класса, выдаваемый классификатором в результате его применения к данному конкретному объекту [4]. Когда говорят о нейросетевой классификации, подразумевают, что в качестве целевого алгоритма (классификатора) выступает нейронная сеть.

Рассмотрим более подробно этапы построения нейросетевого классификатора при решении конкретной прикладной задачи (рис. 1).



**Рис. 1. Этапы построения нейросетевого классификатора**

Предположим, что задача классификации возникла – определена предметная область, известен род объектов, которые требуется классифицировать, определено конечное множество классов.

Первый этап, с которым приходится сталкиваться на практике – это подготовка данных. На этом этапе требуется:

1. собрать выборку объектов,
2. произвести их разметку (то есть получить априорное распределение объектов из выборки на классы),
3. произвести валидацию выборки (т.е. убедиться, что количество объектов каждого класса достаточно, что данные не зашумлены, и что выборка репрезентативна),
4. определить множество признаков, которыми объекты будут описываться и предъявляться модели,
5. перевести объекты в признаковое пространство, получив таким образом матрицу «объект-признак» и окончательно формализовав задачу.

Следующим этапом после подготовки данных является построение модели нейронной сети. На этом этапе требуется выбрать архитектуру сети, определиться с количеством слоев, количеством и типом нейронов в слоях.

Третьим этапом является этап обучения модели. Обучение нейронных сетей чаще всего происходит путем итеративной оптимизации функции ошибки в пространстве настраиваемых параметров модели. На этом этапе требуется задать функцию ошибки, выбрать алгоритм обучения и его параметры, задать критерий останова процедуры обучения и собственно произвести обучение.

После завершения обучения происходит тестирование обученной модели. Для этого производится обработка заранее отложенной независимой тестовой выборки. Должен быть задан критерий качества обработки тестовой выборки. Если качество обработки тестовых данных устраивает разработчика, процесс построения нейросетевого классификатора считается завершенным, и полученная нейронная сеть может применяться для классификации новых объектов.

Выше был описан идеальный сценарий процесса построения нейросетевого классификатора, когда после множества действий, каждое из которых вариативно, разработчик сразу устроило качество полученной модели (на рис. 1 все этапы пройдены «сверху-вниз» по сплошным стрелкам). В реальности, качество полученной модели может не сразу устроить разработчика, и тогда наиболее логично начать возвращаться к предыдущим

этапам постепенно «снизу-вверх» (на рис. 1 по штриховым стрелкам), то есть в первую очередь попробовать изменить процедуру обучения, обучить модель снова, и снова проверить качество на тестовой выборке. Если не удастся добиться приемлемого качества за счет изменения процедуры обучения, приходится возвращаться к этапу построения модели и выбора архитектуры, и, изменив архитектуру, снова обучать ее и проверять качество. Если же не удастся достичь приемлемого качества, перепробовав различные варианты архитектур и параметров процедуры обучения, приходится обращаться к этапу подготовки данных с возможным дополнением выборки или изменением признакового пространства и снова спускаться к этапу проверки качества, подбирая архитектуру и параметры процедуры обучения.

Как видно, на каждом из этапов процесса построения классификатора приходится пробовать различные значения гиперпараметров, большинство из которых несет в себе вариативность. Кроме того, проблема заключается еще и в том, что между этапом выбора архитектуры и пониманием, позволяет данная архитектура решать поставленную задачу или нет, стоит этап обучения, который является ресурсозатратным, и его приходится повторять множество раз.

Этап сбора и подготовки данных слабо поддается формализации и автоматизации, так как он во многом опирается на экспертные знания: для разметки выборки и определения множества признаков чаще всего привлекаются эксперты из предметной области.

Этап обучения, хотя и является ресурсозатратным, но тем не менее, большинство алгоритмов обучения нейросетей достаточно универсальны, для них существуют рекомендуемые параметры, а также разработаны библиотеки автоматизированной оптимизации гиперпараметров [5, 6, 7].

Наибольшая вариативность кроется на этапе выбора архитектуры модели. При этом кроме описанных выше существующих общих идей и соображений по такому выбору, в литературе не было найдено исследований или рекомендаций по оценке сложности требуемой архитектуры в зависимости от свойств обучающей выборки. Представляется перспективным разработка методов оценивания сложности архитектуры нейронной сети, требуемой для решения конкретной формализованной задачи классификации.

### Архитектуры нейросетевых классификаторов

Базовой нейросетевой архитектурой для решения задачи классификации можно считать многослойный перцептрон. Многослойный перцептрон вычисляет выходной вектор по входному  $x^{(N)} = f(x^{(0)})$ , при этом реализует алгебраическое преобразование вида:

$$x_i^{(k)} = s \left( \sum_{j=1}^{m_{k-1}} w_{ij}^{(k)} x_j^{(k-1)} \right), \quad k = \overline{1, N}, \quad i = \overline{1, m_k}$$

где  $N$  – количество слоев нейросети,  $x^{(k)} \in \mathbb{R}^{m_k}$  – вектор выходов слоя  $k$ ,  $x^{(0)} \in \mathbb{R}^{m_0}$  – вектор входов нейросети,  $m_k$  – количество нейронов слоя  $k$ ,  $m_0$  – количество входов первого слоя нейросети,  $w_{ij}^{(k)} \in \mathbb{R}$  – весовой коэффициент  $j$ -го входа  $i$ -го нейрона слоя  $k$ ,  $s(\cdot)$  – дифференцируемая функция, называемая функцией активации.

В 1987 году Хехт-Нильсен на основе работ А.Н. Колмогорова доказал представимость непрерывной функции многих переменных с помощью двухслойной нейросети [8]. Хотя стоит признать, что этот результат неконструктивен, тем не менее, он дает основание считать многослойный перцептрон универсальным аппроксиматором функций многих переменных.

Единственной оценкой сложности многослойного перцептрона в зависимости от обучающей выборки, встречающейся в литературе, является утверждение, что количество параметров (весовых коэффициентов) нейросети не должно превышать объем обучающей

выборки, иначе модель будет стремиться запомнить все обучающие примеры, не стремясь к обобщению[9]. Однако современные методы обучения, на практике доказавшие свою применимость, используют различные методы регуляризации обучения (dropout, батч-нормализация), за счет чего становится возможным хорошее обучение моделей с количеством весов многократно превышающим объемы выборок[10].

Бóльших успехов в части развития архитектур, их сравнения и выбора, исследователи добились при обработке изображений с помощью сверточных нейронных сетей. Идеи сверточных нейросетей были вдохновлены поведением человеческого зрительного аппарата, и впервые были затронуты в когнитроне Фукушимы. Пионерской работой с области сверточных сетей считается работа Яна Ле Куна [11]. На основе описанных в работе [11] принципов затем родилось множество модификаций и вариаций сверточных сетей, таких как AlexNet, VGG, GoogLeNet и др. Архитектурным отличием сверточных сетей от многослойного персептрона можно считать ограниченную рецептивную область нейронов и существенно большее количество последовательных слоев. Здесь особо можно выделить семейство архитектур ResNet, число слоев в которых достигает сотен и даже тысяч. Именно с развитием и появлением сверточных нейросетей, состоящих из десятков слоев, связано появление термина «глубокое обучение». Подробный обзор архитектур сверточных нейронных сетей можно найти в [12].

Отличительной особенностью задачи классификации изображений является то, что для таких данных не требуется разрабатывать признаковое пространство, так как сами пиксели исходного изображения выступают в качестве входных признаков. Предобработка данных заключается лишь в приведении различных изображений к единому размеру. Вместе с этим, стоит отметить, что данная особенность несет за собой сложно поддающуюся осмыслению размерность признакового пространства, фактически соответствующую разрешению обрабатываемых изображений.

Рассмотрев архитектуры различных сверточных нейронных сетей, можно отметить, что насколько бы не была сложна и глубока архитектура сверточных слоев, в конечном итоге на выходе сети расположены полносвязные слои, то есть, фактически, многослойный персептрон. Можно сказать, что сверточные слои являются сложным нейросетевым предобработчиком, выделяющим наиболее информативные признаки из входных данных, которые затем подаются на вход многослойного персептрона. В исходном предположении о том, что входные данные имеют пространственную структуру, применение сверточных слоев позволяет эффективно выделять информативные признаки.

### **Методы синтеза архитектур нейросетевых классификаторов**

Существует два подхода по направленному синтезу архитектур нейронных сетей. Один из них носит название Brain Damage (от англ. «повреждение мозга») [13] и основывается на том, что в качестве начальной архитектуры выбирается перепараметризованная модель, способная решать целевую задачу. Происходит первичное обучение такой модели и оценка качества ее работы. Затем в обученной модели происходит поиск незначимых параметров (весовых коэффициентов), которые вносят малый или незначительный вклад в решение целевой задачи. Незначимые параметры принимаются равными нулю и происходит повторное обучение и оценка работы такой «урезанной» модели. Процесс итеративно повторяется до тех пор, пока не будет найден компромисс между сложностью модели (количеством ненулевых весовых коэффициентов) и качеством решения задачи.

Другим подходом является так называемый Neural Architecture Search [14, 15, 16]. Данный подход напротив, начинается с того, что выбирается некоторая простая начальная архитектура нейронной сети, которая затем итеративно наращивается различными блоками нейронов. На каждой итерации синтеза происходит генерация нескольких вариантов архитектуры, усложненных по сравнению с предыдущей, обучение каждого из вариантов, и выбор наилучшего по заданному критерию качества. Процесс продолжается, пока на

очередной итерации не прекратится значимое улучшение значения критерия качества или не будет достигнут заданный предел сложности результирующей модели. Обзор работ, посвященных Neural Architecture Search, можно найти в [17].

Оба описанных подхода могут применяться на этапе выбора архитектуры нейронной сети (этап 2 на рис. 1) и позволяют целенаправленно изменять архитектуру для достижения желаемого качества решения задачи. Однако, как уже отмечалось ранее, после внесения изменений в архитектуру модели ее неизбежно приходится обучать заново, и повторять эти действия многократно, затрачивая большое количество времени и вычислительных ресурсов.

Практический опыт использования глубоких сверточных нейронных сетей показывает, что при прочих равных лучшее качество дают более глубокие архитектуры с большим количеством параметров. В последние годы появились результаты, показывающие, что на перепараметризованных моделях наблюдаются интересные эффекты «двойного спуска», как по сложности модели, так и по количеству итераций обучения, позволяющие получать более качественные решения [18, 19]. Возникает желание использовать такой мета-подход к выбору архитектуры, при котором задается заранее перепараметризованная модель с расчетом на то, что она будет оптимизирована в процессе обучения. Такие тенденции, опять же, не говорят о рациональности выбора архитектуры. К тому же, такой подход находит применение лишь в сверточных нейронных сетях при классификации изображений или сводимых к таковым. В прочих задачах классификации, где рассчитывать на пространственную структуру данных сложно, и поэтому сложно рассчитывать на применение сверточных сетей, остается использовать многослойный перцептрон, эмпирически подбирая количество слоев и нейронов.

### Заключение

В статье была затронута задача синтеза архитектур нейронных сетей при решении задачи классификации. Был рассмотрен процесс построения нейросетевых классификаторов, в котором этап выбора архитектуры несет в себе большую вариативность, из-за чего задача может становиться чрезмерно ресурсозатратной. При этом в ряде задач может возникать потребность в «минимальных» архитектурах. Развитые архитектурные решения применяются в области классификации изображений сверточными нейронными сетями, но в случаях, когда нельзя рассчитывать на пространственную структуру данных, приходится обращаться к классической архитектуре многослойного перцептрона. В большинстве случаев обоснование выбора архитектуры носит эмпирический характер, используется метод аналогии. В связи с этим возникает желание построения оценок на минимальную архитектуру ИНС хотя бы в некотором классе задач и архитектур. В качестве первого приближения к решению может быть выбрана задача бинарной классификации многослойным перцептроном в пространстве признаков невысокой размерности (не более 10). Это позволит исследовать применимость размерности Вапника-Червоненкиса для экспериментальной оценки сложности нейронных сетей.

### Библиографический список

1. **Emmert-Streib, F., Yang, Z., Feng, H., Tripathi, S., Dehmer, M.** An Introductory Review of Deep Learning for Prediction Models With Big Data // *Frontiers in Artificial Intelligence*, vol. 3, 2020. Available at: <https://www.frontiersin.org/article/10.3389/frai.2020.00004>.
2. **Eliseev, Vladimir L., Miliukova, Ekaterina A., Kolpinskiy, Sergey V.** Neural Network Cryptographic Obfuscation for Trusted Cloud Computing // *CEUR Workshop Proceedings*, vol. 2965, Russian Advances in Fuzzy Systems and Soft Computing: Selected Contributions to the 10th International Conference on "Integrated Models and Soft Computing in Artificial Intelligence", IMSC 2021, Kolomna, Russia, May 17-20, 2021. - p. 201-207. Available at: <http://ceur-ws.org/Vol-2965/paper26.pdf>.



3. **Hastie, T., Tibshirani R., Friedman J.** Chapter 7.9. Vapnik–Chervonenkis Dimension // The Elements of Statistical Learning: Data Mining, Inference, and Prediction. — 2nd ed. — Springer-Verlag, 2009. — 746 p. — ISBN 978-0-387-84857-0.
4. Классификация [Электронный ресурс]: MachineLearning.ru. URL: <http://www.machinelearning.ru/wiki/index.php?title=Классификация> (дата обращения 13.05.2022).
5. **Bergstra, J., Yamins, D., CoxMaking, D. D.** a Science of Model Search: Hyperparameter Optimization in Hundreds of Dimensions for Vision Architectures // Proceedings of the 30<sup>th</sup> International Conference on Machine Learning, Atlanta, Georgia, USA, 2013. JMLR: W&CP vol. 28. Available at: <http://proceedings.mlr.press/v28/bergstra13.pdf>.
6. Hyperas [Электронный ресурс]: Keras + Hyperopt: A very simple wrapper for convenient hyperparameter optimization. URL: <https://maxpumperla.com/hyperas/> (дата обращения 13.05.2022).
7. Hyperopt [Электронный ресурс]: Distributed Asynchronous Hyperparameter Optimization in Python. URL: <https://github.com/hyperopt/hyperopt> (дата обращения 13.05.2022).
8. **Hecht-Nielsen, R.** Kolmogorov’s mapping neural network existence theorem // IEEE First Annual Int. Conf. on Neural Networks, San Diego, 1987. Vol. 3. pp. 11-13.
9. **Хайкин, С.** Нейронные сети: полный курс, 2-е издание. Издательский дом Вильямс, 2008.
10. **Николенко, С., Кадури, А., Архангельская, Е.** Глубокое обучение. – СПб.: Питер, 2019. – 480 с.: ил. – (Серия «Библиотека программиста»). ISBN 978-5-496-02536-2.
11. **Le Cun, Y. et al.** Handwork digit recognition with a back-propagation network // Advances in Neural Information Processing Systems. Colorado, USA: [s. N], 1990: 396-404.
12. **Mohammed, A. T.** Review of deep convolution neural network in image classification // 2017 International Conference on Radar, Antenna, Microwave, Electronics, and Telecommunications. Available at: <https://core.ac.uk/reader/275672004>.
13. **Le Cun, Y., Denker, J., Solla, S.** Optimal Brain Damage // Advances in Neural Information Processing Systems, 1989, vol. 2. Available at: <https://proceedings.neurips.cc/paper/1989/hash/6c9882bbac1c7093bd25041881277658-Abstract.html>.
14. **Zoph, B. and Le, Q. V.** Neural Architecture Search with Reinforcement Learning // arXiv, February 15, 2017. Available at: <http://arxiv.org/abs/1611.01578>.
15. **Elsken, T., Metzen, J.-H. and Hutter, F.** Simple and Efficient Architecture Search for Convolutional Neural Networks // arXiv, November 13, 2017. Available at: <http://arxiv.org/abs/1711.04528>.
16. **Cai, H., Zhu, L. and Han, S.** ProxylessNAS: Direct Neural Architecture Search on Target Task and Hardware // arXiv, February 22, 2019. Available at: <http://arxiv.org/abs/1812.00332>.
17. **Elsken, T., Metzen, J. H. and Hutter, F.** Neural Architecture Search: A Survey // arXiv, April 26, 2019. Available at: <http://arxiv.org/abs/1808.05377>.
18. **Belkin, M., Hsu, D., Ma, S. and Mandal, S.** Reconciling modern machine learning practice and the bias-variance trade-off // arXiv, September 10, 2019. Available at: <http://arxiv.org/abs/1812.11118>.
19. **Nakkiran, P., Kaplun, G., Bansal, Y., Yang, T., Barak, B., Sutskever, I.** Deep Double Descent: Where Bigger Models and More Data Hurt // arXiv, December 4, 2019. Available at: <http://arxiv.org/abs/1912.02292>.

## РАЗРАБОТКА МОДУЛЬНОЙ СИСТЕМЫ МОДЕЛИРОВАНИЯ СЕТИ КВАНТОВОГО РАСПРЕДЕЛЕНИЯ КЛЮЧЕЙ

В.Л. Елисеев

ORCID:0000-0002-9341-7475 e-mail:vlad-eliseev@mail.ru

В.В. Щербаков

ORCID:0000-0002-4327-2976 e-mail:vitya.shcherbakov@gmail.com

Национальный исследовательский университет «Московский энергетический институт»

Современные криптографические алгоритмы, играющие большую роль в технологиях защиты информации, потенциально могут быть скомпрометированы с помощью эффективных квантовых компьютеров, появление которых ожидается в ближайшем будущем. Технология квантового распределения ключей (КРК) лишена этого недостатка и может стать решением сложившейся проблемы. Однако, сети КРК обладают рядом важных особенностей. Так, в частности, такие сети состоят из доверенных промежуточных узлов (ДПУ), участвующих в передаче применяемых для шифрования данных секретных ключей, которые вырабатываются с использованием квантовых эффектов. Генерация ключей на сегментах происходит с переменной интенсивностью, что отражается на производительности и эффективности сети. Определение применимости тех или иных алгоритмов маршрутизации и пропускной способности сегментов и участков сети при ее различных конфигурациях представляют отдельный исследовательский интерес, поэтому для проведения задач исследования работы сети КРК была разработана модульная система моделирования на основе подхода систем массового обслуживания. Данный подход является наиболее близким к реальному применению сети, обеспечивая имитацию потока заявок от шифраторов на обслуживание сетью КРК. Модульность системы заключается в разбиении модели на несколько взаимодействующих частей, моделирующих конкретные аспекты сети — ее топологию, маршрутизацию, вычисление метрик, порядок обслуживания очереди заявок и непосредственно имитирующую пользовательские запросы часть, что позволит изучить влияние того или иного параметра на состояние моделируемой сети.

**Ключевые слова:** квантовое распределение ключей, квантовая сеть, маршрутизация, система моделирования, прикладное программное обеспечение

## DEVELOPMENT OF QUANTUM KEY DISTRIBUTION NETWORK MODELING SOFTWARE

V.L. Eliseev, V.V. Shcherbakov

National research university «Moscow power engineering institute»

**Purpose:** The article describes applicability of queueing theory approach in quantum key distribution (QKD) network modeling software both with possible usage of this software in QKD network researches including quantum link throughput and various routing algorithms efficiency analysis.

**Design / methodology / approach:** The paper explains modular QKD network modeling software structure employing closely tied to it object-oriented programming paradigm. With support of inheritance and polymorphism comes ability to fine-tune the modules, each representing separate QKD network aspect, allowing to flexibly setup various network instances. Aforementioned features are providing suitable conditions for experimenting with usage of modules with distinct policies, including queueing, request generating imitation, network topology, quantum key generation models and pathfinding algorithms.

**Findings:** Provided software product is capable for QKD network simulation and can be used in scientific and engineering researches or for educational purposes.

**Research limitations/implications:** Presented article is aimed to be a starting point for more detailed and thorough research works in field of QKD networking simulation.

**Originality/value:** QKD networks are considered to be a possible answer to approaching quantum computer advent, which may render many of classic cryptographic ciphers unreliable. Further researches in this field will follow, and presented QKD modeling tool is the first of such kind.

**Keywords:** quantum key distribution, quantum network, routing, modeling system, software application

## Введение

В современном мире всё большее значение для компьютеров и сетей передачи данных приобретают различные технологии информационной безопасности и, в частности, одна из важнейших – криптография. Криптографические методы защиты информации позволяют обеспечить конфиденциальность и целостность данных. Симметричные алгоритмы шифрования с помощью секретного ключа позволяют преобразовывать исходную информацию в шифротекст и обратно. Секретные ключи являются неотъемлемой частью криптографических протоколов, поскольку надежность криптографической системы во многом зависит именно от секретности ключа, а не алгоритма шифрования.

Для обмена зашифрованными сообщениями обе стороны должны обладать общим секретным ключом, который будут применять для зашифрования и расшифрования сообщений. Применение симметричных блочных шифров требует периодической смены ключа шифрования, поскольку использование одного и того же ключа долгое время в итоге снижает стойкость криптосистемы. Именно для предотвращения снижения стойкости криптосистемы применяют протоколы распределения ключей [1].

Протокол Диффи-Хеллмана – криптографический протокол, позволяющий двум сторонам получить общий секретный ключ, используя незащищенный от прослушивания канал связи. Этот протокол является первым общеизвестным и одним из самых распространенных протоколов распределения ключей. Принцип его работы основан на высокой вычислительной сложности дискретного логарифмирования больших чисел, не позволяющей взломать протокол на современных вычислительных машинах.

В последние годы активно развиваются так называемые квантовые алгоритмы и технологии квантовых компьютеров. Отчасти это обусловлено привлекательными возможностями по ускорению вычислений по сравнению с классическими компьютерами в некоторых частных случаях, к которым относится задача дискретного логарифмирования. Квантовый алгоритм Шора позволяет решить задачу дискретного логарифмирования, а также сводимую к ней задачу факторизации чисел на простые множители за полиномиальное время, что приводит к недопустимому снижению криптостойкости протокола Диффи-Хеллмана. Таким образом, вероятное создание мощного и эффективного квантового компьютера в ближайшем будущем может привести к компрометации целого ряда современных криптографических алгоритмов и протоколов, включая RSA, HTTPS, и т.д. [2].

### Квантовое распределение ключей

Одним из способов решения описанной проблемы может стать квантовое распределение ключей (КРК). В основе КРК лежит идея об использовании квантовых состояний для передачи информации, на основе которой возможно создать общие секретные ключи у двух абонентов. Квантовые состояния не могут быть в точности повторены, а любая попытка их измерить приведет к их изменению. Секретность вырабатываемых квантовых ключей обусловлена не сложностью решения какой-либо вычислительной задачи, а законами физики.

В качестве носителя квантового состояния выступает передаваемый по оптоволоконным линиям связи одиночный фотон, модулированный особым образом – по поляризации, фазе или иным способом. Передача одиночных фотонов на большие расстояния сопряжена с рядом объективных трудностей, прежде всего – с возрастающим влиянием помех при распространении в плотной среде. Каким-либо образом увеличить соотношение сигнал/шум не предоставляется возможным, поскольку это противоречит самому принципу метода (передача и регистрация одиночных фотонов). На данный момент максимальное расстояние по оптоволокну, на котором можно рассчитывать на выработку квантовых ключей, составляет примерно 100 км [3].

Выработанный квантовый ключ можно использовать в сетевых шифраторах для обеспечения конфиденциальности и целостности передаваемых данных [4].

Очевидно, что дальности распределения секретных ключей в 100 км недостаточно для нужд современных сетей. Проектирование сетей КРК по принципу доверенных промежуточных узлов (ДПУ) позволяет преодолеть данное ограничение. Вместе с понятием квантового ключа (КК) – ключа, выработанного на основе квантовых состояний, вводится понятие квантово-защищенного ключа (КЗК) – классического секретного ключа, зашифрованного с помощью КК и, тем самым, сопоставимого с ним по своей стойкости. Наиболее просто такое сопоставление может быть сделано при шифровании КЗК с помощью КК на основе одноразового блокнота. При этом, согласно принципам корректного применения одноразового блокнота [1], КК после однократного использования должен уничтожаться. На каждом сегменте сети, участвующем в передаче, КЗК перешифровывается новым КК, вырабатываемым смежными сегменту узлами, тем самым, КЗК из одного узла может быть доставлен через доверенные узлы-посредники в любой другой, не обязательно лежащий в пределах максимальной дальности передачи квантовых состояний. Очевидно, что такая сеть хорошо масштабируется, поскольку максимальное расстояние между крайними узлами сети может быть произвольным и ограничено только количеством промежуточных узлов между ними. КЗК можно использовать в сетевых шифраторах, находящихся в крайних узлах сети.

### **Маршрутизация в сетях квантового распределения ключей**

В распределенных сетях КРК особую роль играет маршрутизация – определение цепочки узлов, через которые будет осуществляться передача КЗК между двумя целевыми узлами. В силу своих физических свойств выработка КК происходит с переменной интенсивностью. В то же время, КЗК при передаче по сети расходуют КК на сегментах. Поэтому два процесса происходят в сети одновременно: выработка новых КК на всех сегментах сети и их потребление на сегментах сети, через которые проходят КЗК. Очевидно, что эти процессы приводят к определенной динамике присутствия доступных для применения КК на сегментах. Сегменты, на которых есть выработанные КК или вырабатывающие КК быстрее, чем другие, являются предпочтительными для включения их в маршрут передачи КЗК. Чтобы повысить производительность работы сети, целесообразно организовать КК в автоматическом режиме и их сохранение в специальных буферах для дальнейшего использования. Таким образом, возникает задача построения маршрута для передачи КРК с целью оптимального использования запаса КК на сегментах. Известным научным результатом в области маршрутизации является алгоритм Дейкстры поиска кратчайшего пути в графе [5], на основе которого разработаны многие алгоритмы и протоколы маршрутизации в классических сетях передачи данных.

Алгоритм Дейкстры для взвешенного неориентированного графа предназначается для поиска кратчайших путей от какой-либо заданной вершины до всех других в графе. Кратчайшим называется путь, сумма весов ребер которого минимальна (вес в данном случае выступает подобием «расстояния») [6]. В качестве метрики (весов вершин) в случае сетей КРК имеет смысл использовать количество квантовых ключей, сохраненных в буферах узлов, либо величину, обратную интенсивности генерации квантовых ключей в момент времени. Далее в настоящей работе будем в качестве метрики использовать количество КК в буферах промежуточных узлов.

### **Квантовые сети как системы массового обслуживания**

Рассмотрим характерные особенности передачи КЗК по запросу пары сетевых шифраторов, находящихся в различных узлах квантовой сети. Как уже отмечалось, для передачи КЗК в таком случае необходимо построить маршрут, выбрав наиболее предпочтительные сопряженные друг с другом сегменты квантовой сети. В сложной сети с множественными путями из одного узла сети в

другой последовательные запросы могут приводить к построению разных маршрутов, что обусловлено динамикой процесса выработки и потребления КК. При этом, в разветвленной сети из многих узлов и сетевых шифраторов запросы КЗК могут происходить в разных парах узлов, что также будет влиять на динамику построения маршрутов.

Естественным образом возникает задача исследования пропускной способности квантовой сети в части удовлетворения запросов КЗК со стороны шифраторов. Данная задача может рассматриваться в терминах теории массового обслуживания, при этом квантовая сеть представляет собой сложное устройство в системе массового обслуживания (СМО), в которое поступают заявки на выдачу секретного ключа для пары шифраторов. Обслуживание клиентов сети происходит посредством очереди заявок. В системе определен порядок обслуживания заявок, емкость очереди заявок, а также срок действия заявки. Отказы на обслуживание производятся при невозможности построения пути между запрошенными узлами, либо при истечении срока жизни заявки [7]. Интенсивность выработки квантовых ключей ограничена, что отражается на производительности сети. С ростом масштабов квантовых сетей неизбежно придет потребность в моделировании сетей с точки зрения качества обслуживания – количества удовлетворенных запросов на выдачу квантовых ключей. В данных условиях имеет значения задача моделирования пропускной способности квантовых сетей. Анализ задачи на основе изложенной выше информации показывает, что пропускная способность квантовой сети как СМО зависит от: топологии сети; производительности выработки КК на сегментах; ёмкости буферов КК на узлах сети; расположения шифраторов и частоты запросов КЗК от них. Учитывая сложность комплексной задачи оценки пропускной способности квантовой сети и новизну темы построения распределенных квантовых сетей, представляется целесообразным разработать систему для моделирования квантовой сети с возможностью задания произвольной топологии сети и параметров, характеризующих её работу. В качестве примера такой сети рассмотрим квантовую сеть Москва – Санкт-Петербург протяженностью 700 км, запущенную в эксплуатацию в 2021 году. Эта сеть должна стать началом разветвленной сети протяженностью 7000 км, которая должна быть построена к 2024 г. [8].

### Система моделирования

Для исследования задачи маршрутизации в сети КРК была разработана модульная система моделирования по принципу систем массового обслуживания для имитации потока заявок от шифраторов и их обслуживания сетью КРК.

Модульность системы заключается в разбиении модели на несколько взаимодействующих программных частей, обслуживающих конкретные аспекты сети – её топологию, маршрутизацию, имитацию выработки ключей, порядок обслуживания очереди и непосредственно имитирующую пользовательские запросы часть, что позволит более тесно изучить влияние того или иного параметра на конечное состояние моделируемой сети.

Структура системы моделирования представлена на рис. 1.

Модуль *Topology* хранит информацию о топологии сети в виде односвязного неорграфа.

Модуль *Request Generator* производит имитацию поступления заявок на обслуживание парами шифраторов узлов, выбирая их, обращаясь к модулю *Topology*. Сформированные заявки содержат время поступления и истечения срока действия, а также приоритет обслуживания.

Модуль *Queue Manager* организует очередь выработанных модулем *RequestGenerator* заявок, определяет порядок обслуживания очереди (FIFO, LIFO, очередь с приоритетом) и её емкость, принимая во внимание различные параметры (вышеуказанные время истечения срока жизни, приоритет и возможные другие).

Модуль *Pathfinder* извлекает заявки из очереди *QueueManager*, задействует выбранный алгоритм маршрутизации, используя информацию о сети модуля *Topology*, возвращает найденный

путь, либо сигнализирует о невозможности построения пути в связи с отсутствием квантовых ключей на сегментах или физического соединения между узлами.

Модуль *Key Generation Model* осуществляет имитацию выработки и расходования квантовых ключей на сегментах и организует хранение выработанных ключей. Механизм имитации представляет собой случайное изменение интенсивности выработки по выбранному вероятностному закону с заданными параметрами.

Центральный модуль *QKD\_Network* организует взаимодействие модулей друг с другом и управляет системой в целом.

Модульность системы позволяет компоновать друг с другом различные модули, получая различные варианты сетей. Система моделирования реализована на языке C++ и работает под операционной системой Linux.

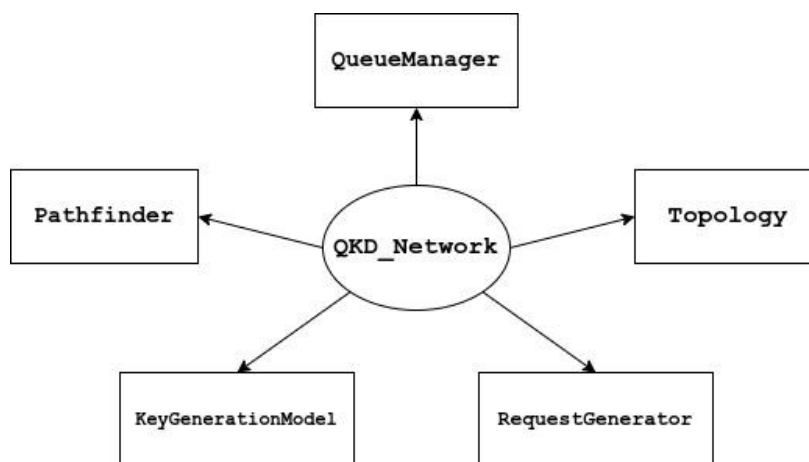


Рис. 1. Структура системы моделирования

### Результат моделирования

В качестве объекта моделирования была выбрана запущенная в 2021 г. квантовая сеть Москва – Санкт-Петербург:

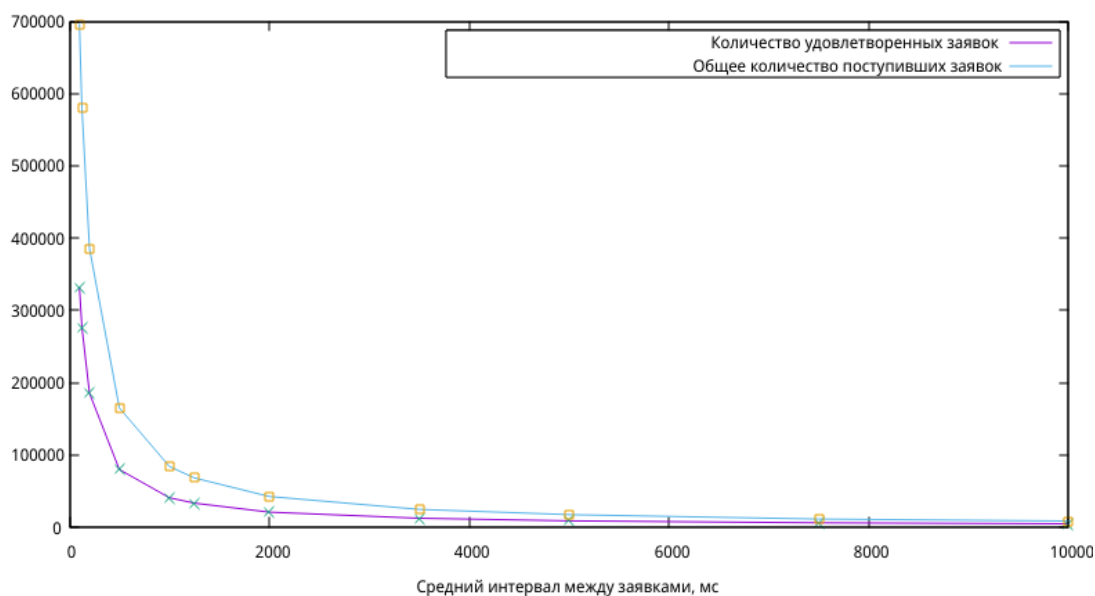


Рис. 2. Квантовая сеть Москва - Санкт-Петербург

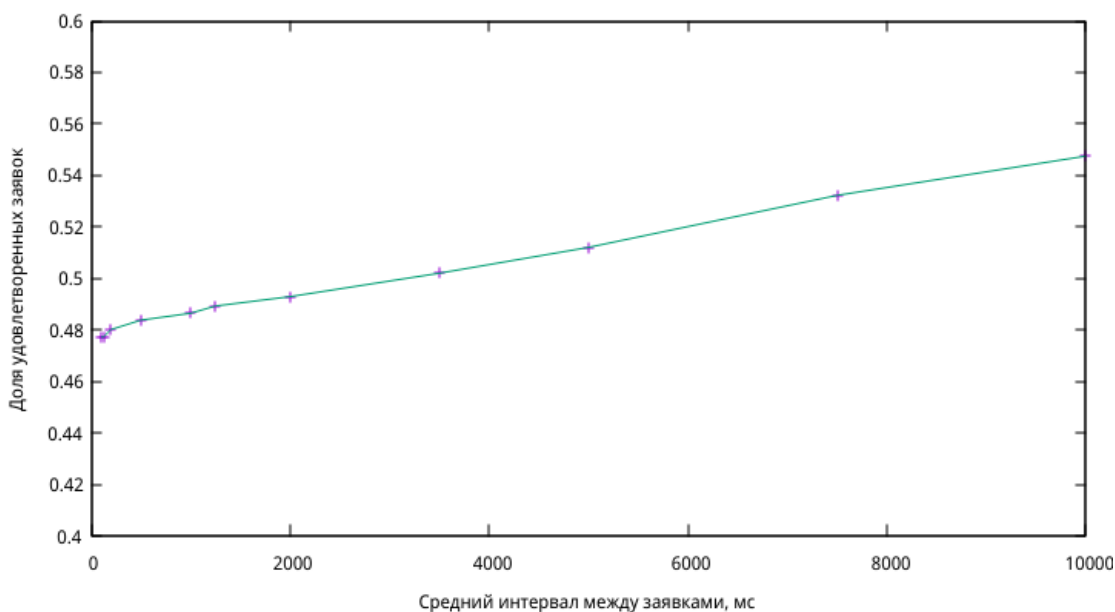
Было проведено моделирование работы данной сети для определения зависимости ее пропускной способности от интенсивности поступления в нее заявок на обслуживание. Заявки на выдачу ключа поступают для целевых узлов (обозначены зеленым цветом на рис. 2), остальные узлы являются промежуточными. Параметры проведенного моделирования:

- тип СМО согласно нотации Кендалла – М/М/1 (время между поступлениями заявок и время их обслуживания подчиняются пуассоновскому закону, количество одновременно обслуживаемых заявок – 1);
- интенсивность выработки квантовых ключей (бит/с) подчиняется экспоненциальному закону с параметром  $\lambda = 0.005$ ;
- размер очереди заявок – 10, порядок обслуживания – FIFO («первым пришел - первым ушел»);
- время моделирования – 24 часа.

Полученные зависимости общего количества заявок и количества обслуженных заявок (рис. 3) и доли обработанных заявок от среднего значения интервала между заявками (рис. 4):



**Рис. 3. Зависимости общего количества заявок и количества обслуженных заявок от среднего интервала времени между заявками**



**Рис. 4. Зависимость доли обслуженных заявок от среднего интервала времени между заявками**



По полученным графикам (рис. 3 и 4) видно, что при заданных параметрах пропускная способность сети заданной сети КРК выше при бóльшем временном интервале между заявками, а зависимости количества заявок от временного интервала между прибытиями в систему заявок носят экспоненциальный характер.

### Заключение

Была разработана система имитационного моделирования сетей квантового распределения ключей, а также проведено моделирование работы квантовой сети Москва-Санкт-Петербург с целью выяснения ее пропускной способности. В дальнейшем планируется усовершенствовать систему, добавив более реалистичную модель выработки и потребления квантовых ключей. Разработанная система позволяет производить моделирование сетей произвольной топологии, а также задавать различные параметры работы элементов сети.

### Библиографический список

1. **Шнайер Б.** Прикладная криптография. Протоколы, алгоритмы, исходные тексты на языке Си – М.: Триумф, 2002. – 816 с.
2. **Граймс Р. А.** Апокалипсис криптографии. Пер. с англ. В. А. Яроцкого. М.: ДМК Пресс, 2020. – 290 с.
3. **M. Mehic, M. Niemiec, S. Rass, J. Ma, M. Deev, A. Aguado, V. Martin, S. Schauer, A. Poppe, C. Pacher, M. Voznak** Quantum Key Distribution: A Networking Perspective // ACM Comput. Surv. 53, 5, Article 96 (September 2020). 41 pages.
4. **Втюрина А.Г., Елисеев В.Л., Жилиев А.Е., Николаева А.С., Сергеев В.Н., Уривский А.В.** Реализация средства криптографической защиты информации, использующего квантовое распределение ключей // Доклады Томского государственного университета систем управления и радиоэлектроники. – 2018. – Т. 21. – № 2. – С. 15-21. – DOI 10.21293/1818-0442-2018-21-2-15-21.
5. **Tsai, C.-W.; Yang, C.-W.; Lin, J.; Chang, Y.-C.; Chang, R.-S.** Quantum Key Distribution Networks: Challenges and Future Research Issues in Security. Appl. Sci. 2021, 11, 3767.
6. **Кормен, Томас Х. и др.** Алгоритмы: построение и анализ, 3-е изд. : Пер. с англ. – М. : ООО «И.Д. Вильямс», 2013. – 1328 с. : ил. – Парал. тит. англ.
7. **Солнышкина, И.В.** Теория систем массового обслуживания : учеб. Пособие / И. В. Солнышкина. – Комсомольск-на-Амуре : ФГБОУ ВПО «КнАГТУ», 2015. – 76 с.
8. Первая в РФ линия квантовой связи запущена между Москвой и Петербургом – URL: <https://www.interfax.ru/russia/771209> (дата обращения - 08.05.2022).



**ОПТИМИЗАЦИЯ МОДЕЛИ ПОТЕНЦИАЛЬНЫХ КОНТЕЙНЕРОВ ДЛЯ ПОВЫШЕНИЯ СКОРОСТИ РАЗМЕЩЕНИЯ ОРТОГОНАЛЬНЫХ МНОГОГРАННИКОВ****В.А. Чеканин**

ORCID: 0000-0001-9448-0583 e-mail: vladchekanin@rambler.ru

ФГБОУ ВО «МГТУ «СТАНКИН»

Использование модели потенциальных контейнеров при решении задач компоновки объектов нерегулярной формы, представленных ортогональными многогранниками, обеспечивает полное описание пространства упаковки, что позволяет выбирать наиболее подходящую область контейнера для каждого размещаемого объекта. В статье предложены алгоритмы, решающие задачу сокращения числа обрабатываемых потенциальных контейнеров при размещении ортогональных многогранников, применение которых обеспечивает качественное повышение скорости геометрического конструирования пространства упаковки. Представлены результаты проведённых вычислительных экспериментов на задачах упаковки трёхмерных объектов нерегулярной формы, подтверждающие эффективность разработанных алгоритмов. Предложенная оптимизация модели потенциальных контейнеров делает возможным повышение степени детализации размещаемых ортогональных многогранников для получения более плотной компоновки в разумное время.

**Ключевые слова:** компоновка объектов нерегулярной формы, задача упаковки, модель потенциальных контейнеров, ортогональный многогранник

**OPTIMIZATION OF THE MODEL OF POTENTIAL CONTAINERS TO INCREASE THE SPEED OF PLACEMENT OF ORTHOGONAL POLYHEDRONS****V.A. Chekanin**

Moscow state university of technology «STANKIN»

**Purpose:** Development of algorithms that provide a qualitative increase in the speed of placement of orthogonal polyhedrons when using the model of potential containers to describe placement schemes.

**Design / methodology / approach:** The proposed algorithms are based on reducing the number of potential containers processed when placing objects of complex shape represented by orthogonal polyhedrons. An optimized algorithm for updating a set of potential containers describing the free space of the package solves the problem of determining the areas that will remain unchanged in the container after placing an orthogonal polyhedron, based on the intersection operation applied to the original set of potential containers and the set of potential containers formed in an empty container after placing the considered object. An additional increase in the speed of the model of potential containers is achieved by removing potential containers that cannot be used to place new objects in all possible orientations.

**Findings:** The developed algorithms make it possible to significantly increase the speed of placement of orthogonal polyhedrons. It is shown that with an increase in the number of objects placed using the proposed algorithms, the time spent on placing one object is reduced.

**Research limitations/implications:** Optimization of the model of potential containers makes it possible to place orthogonal polyhedrons specified with a higher degree of detail in a reasonable time to obtain a denser layout of objects.

**Originality/value:** The algorithms proposed in the article make it possible to qualitatively increase the speed of placement of geometrically complex objects when solving a number of time-consuming packaging problems, in particular, designing the layout of aircraft equipment, generating optimized placement of parts on a 3D printer platform.

**Key words:** irregular layout problem, packing problem, model of potential containers, orthogonal polyhedron.

## **Введение**

Задачи компоновки объектов нерегулярной формы являются чрезвычайно трудными и требуют затрат значительных вычислительных ресурсов в ходе их решения, что объясняется как принадлежностью этих задач NP-трудным оптимизационным задачам [1, 2], так и необходимостью детализированного описания геометрически сложных объектов [3]. Эти задачи актуальны при решении большого числа оптимизационных задач нерегулярного раскроя и упаковки, включая задачи проектирования схем размещения деталей на платформе 3D-принтера, фигурного раскроя заготовок промышленных материалов, моделирования микроструктур композитных материалов [4], компоновки летательных аппаратов [5] и технологического оборудования [6].

Можно выделить два направления развития методов повышения скорости решения задач компоновки объектов произвольной геометрии. Первое направление связано с разработкой и развитием эвристических и метаэвристических методов оптимизации, обеспечивающих сокращение числа рассматриваемых альтернативных вариантов решений рассматриваемой задачи [7]. Настоящая работа посвящена второму направлению, связанному с разработкой методов геометрического конструирования пространства упаковки на основе некоторой заранее заданной последовательности размещения объектов.

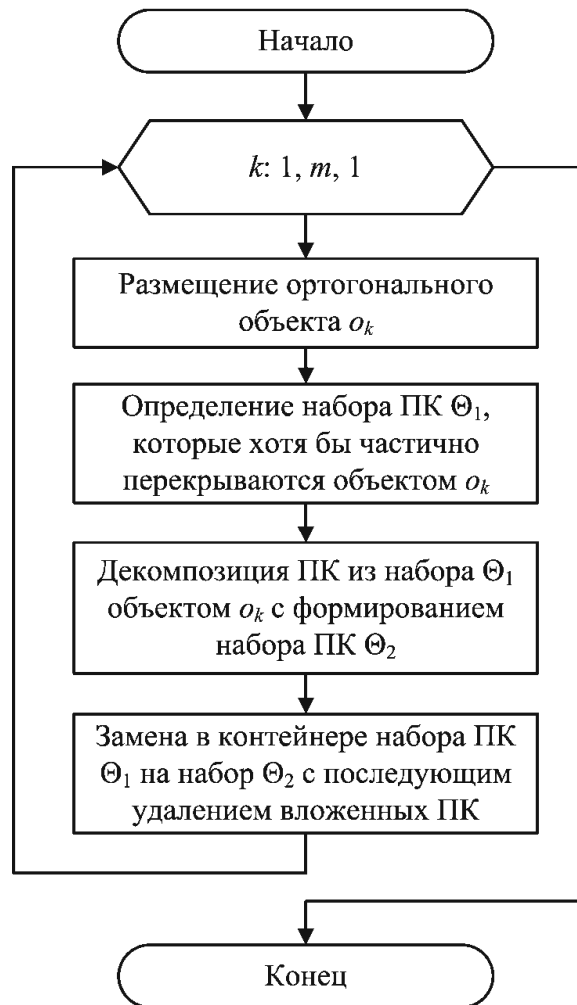
Наиболее точным методом геометрического конструирования упаковки объектов нерегулярной формы является метод, основанный на построении годографа вектор-функции плотного размещения [8-10]. Основным недостатком этого метода является необходимость применения трудоёмких методов нелинейного программирования. Предложенный в статье [11] метод обеспечивает возможность быстрого формирования упаковки объектов сложной формы, представленных в виде ортогональных многогранников (ОМ) [12], полученных в результате вокселизации размещаемых объектов [13-15]. Этот метод основан на использовании модели потенциальных контейнеров (ПК), которая описывает свободное пространство внутри упаковки в виде набора ортогональных областей, предназначенных для размещения новых объектов [7]. В настоящей работе представлены алгоритмы, обеспечивающие повышение скорости размещения ортогональных многогранников.

## **Методика**

Для корректного описания всех свободных пространств внутри упаковки после размещения каждого ОМ необходимо выполнять процедуру обновления набора ПК. Она заключается в получении множества ПК минимальной мощности, которое описывает все существующие свободные области внутри упаковки при условии, что ни один ПК из данного набора даже частично не перекрывает ни один из размещённых объектов. Очевидно, что с увеличением числа ортогональных объектов, входящих в состав размещаемого ОМ, будет расти время, затрачиваемое на обновление набора ПК, поэтому для задач компоновки объектов сложной геометрической формы необходимо применение наиболее быстрого алгоритма обновления набора ПК.

### **Алгоритм последовательного обновления набора ПК (алгоритм 1)**

Первый разработанный алгоритм обновления набора ПК при размещении ОМ заключается в последовательном обновлении ПК после размещения каждого ортогонального объекта, входящего в состав рассматриваемого ОМ (рис. 1).



**Рис. 1. Обновление набора ПК при последовательном размещении всех ортогональных объектов, образующих ОМ**

При создании набора ПК  $\Theta_1$  для каждого ПК  $P_i \in \Theta_1$  должно выполняться условие его перекрытия с ортогональным объектом  $o_k$  :

$$\left(x_i^d < X^d + z_k^d + w_k^d\right) \wedge \left(x_i^d + p_i^d > X^d + z_k^d\right) \quad \forall d \in \{1, \dots, D\},$$

где  $\{p_i^1, p_i^2, \dots, p_i^D\}$  – габаритные размеры ПК  $P_i$ ;  $\{x_i^1, x_i^2, \dots, x_i^D\}$  – координаты точки размещения ПК  $P_i$ .

Набор  $\Theta_2$  будет содержать ПК, полученные в результате декомпозиции ортогональным объектом  $o_k$  всех ПК  $P_i \in \Theta_1$  :

- 1) ПК, сформированные при выполнении условия перекрытия  $\left(X^d + z_k^d > x_i^d\right) \wedge \left(X^d + z_k^d < x_i^d + p_i^d\right) \quad \forall d \in \{1, \dots, D\}$ , которые расположены в точке  $\{x_i^1, x_i^2, \dots, x_i^D\}$  контейнера и имеют габаритные размеры  $\{p_i^1, p_i^2, \dots, p_i^{d-1}, X^d + z_k^d - x_i^d, p_i^{d+1}, \dots, p_i^D\}$ ;
- 2) ПК, сформированные при выполнении условия перекрытия  $\left(X^d + z_k^d + w_k^d > x_i^d\right) \wedge \left(X^d + z_k^d + w_k^d < x_i^d + p_i^d\right) \quad \forall d \in \{1, \dots, D\}$ , которые расположены в

точках  $\{x_i^1, x_i^2, \dots, x_i^{d-1}, X^d + z_k^d + w_k^d, x_i^{d+1}, \dots, x_i^D\}$  контейнера и характеризуются габаритными размерами  $\{p_i^1, p_i^2, \dots, p_i^{d-1}, x_i^d + p_i^d - X^d - z_k^d - w_k^d, p_i^{d+1}, \dots, p_i^D\}$ .

При декомпозиции ПК, перекрывааемых размещаемым ортогональным объектом, могут быть сформированы новые ПК, целиком находящиеся внутри других ПК, что при последующем размещении других объектов приводит к резкому увеличению числа ПК и соответствующему повышению времени, затрачиваемого на формирование упаковки. Поэтому для получения множества ПК минимальной мощности при размещении каждого ортогонального объекта необходимо выполнять поиск и удаление ПК, вложенных друг в друга.

ПК  $P_1$  является вложенным в ПК  $P_2$  (т.е. содержится целиком внутри ПК  $P_2$ ), если выполняется условие:

$$(x_1^d \geq x_2^d) \wedge (x_1^d + p_1^d \leq x_2^d + p_2^d) \quad \forall d \in \{1, \dots, D\}. \quad (1)$$

На первом этапе работы алгоритма поиска вложенных ПК формируется набор  $\Theta'$ , состоящий из ПК  $P_i \in \Theta'$ , для которых  $\exists d \in \{1, \dots, D\}: x_i^d \leq X^d + z_k^d + w_k^d$ . Далее для каждой пары ПК  $P_1, P_2 \in \Theta'$ ,  $P_1 \neq P_2$  выполняется проверка условия (1), при выполнении которого ПК  $P_1$  удаляется из набора, используемого для описания свободного пространства упаковки.

**Алгоритм обновления набора ПК на основе применения операции пересечения (алгоритм 2)**

Второй разработанный алгоритм решает задачу определения областей, которые останутся неизменными в контейнере после размещения ОМ на основе применения к исходному набору ПК  $\Omega_0$  операции пересечения с набором ПК, сформированном в пустом контейнере после размещения рассматриваемого ОМ. Этот алгоритм включает шаги 1–4.

**Шаг 1.** Сформировать набор  $\Omega'_0 \subset \Omega_0$ , состоящий из ПК  $P_i \in \Omega'_0$ , для которых  $\exists d \in \{1, \dots, D\}: x_i^d \leq X^d + S^d$ , где  $S^d = \max(z_k^d + w_k^d)$ ,  $k = 1 \dots |O|$ .

**Шаг 2.** Разместить ОМ  $O$  в точке  $\{X^1, X^2, \dots, X^D\}$  нового пустого контейнера (идентичного по форме и габаритным размерам контейнеру, используемому в решаемой задаче), в результате чего в нём будет сформирован набор ПК  $\Omega_1$ . При размещении ОМ используется описанный выше алгоритм обновления набора ПК на основе последовательного размещения всех ортогональных объектов.

**Шаг 3.** Применить операцию пересечения к наборам ПК  $\Omega'_0$  и  $\Omega_1$ , в результате чего будет получен набор  $\Omega''_0 := \Omega'_0 \cap \Omega_1$ , описывающий свободное пространство исходного контейнера в окрестности размещённого ОМ. При выполнении операции пересечения каждый набор ПК рассматривается в качестве ОМ, состоящего из ортогональных объектов, параметры которых совпадают с параметрами соответствующих ПК.

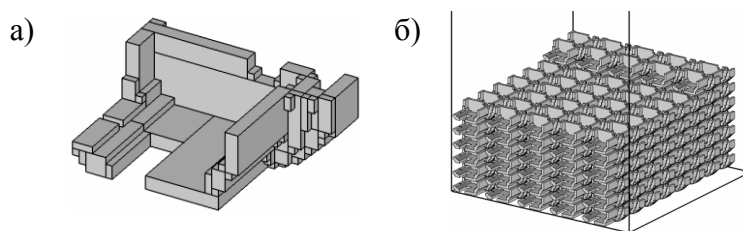
**Шаг 4.** Заменить в наборе  $\Omega_0$  все ПК, входящие в набор  $\Omega'_0$ , на ПК из полученного набора  $\Omega''_0$ .

### Вычислительные эксперименты

Сравнительный анализ эффективности работы разработанных алгоритмов обновления наборов ПК проводился на примере решения задачи размещения набора одинаковых ОМ, имеющих внешние габаритные размеры  $\{62, 42, 24\}$  и состоящих из 66 ортогональных объектов (рис. 2, а). Контейнер представлен параллелепипедом с габаритными размерами

{340, 340, 620}. При размещении ОМ не допускался их поворот (на рис. 2, б представлен пример полученной компоновки набора из 250 ОМ).

Вычислительный эксперимент проводился на персональной ЭВМ (ЦП – Intel Core i5-8400, 2,8 ГГц; ОЗУ – 8 ГБ).



**Рис. 2. Тестовая задача упаковки набора ОМ:**  
а) тип размещаемого ОМ; б) результат размещения набора из 250 ОМ

Сводные результаты решения задачи размещения наборов из 50, 100, 150, 200 и 250 ОМ представлены в табл. 1, в которой используются следующие обозначения:  $T$  – время, затраченное на размещение всех ОМ;  $t$  – среднее время, затрачиваемое на размещение одного ОМ, которое определяется по формуле  $t = T/N$ , где  $N$  – общее число всех ОМ в решаемой задаче. Вычислительный эксперимент показал, что алгоритм 2, основанный на пересечении наборов ПК, обеспечивает качественно более высокую скорость размещения ОМ в сравнении с алгоритмом 1.

**Таблица 1.**

**Тестирование алгоритмов обновления наборов ПК**

Число ОМ в задаче	Число ПК после размещения всех ОМ	Алгоритм 1		Алгоритм 2	
		$T, c$	$t, c$	$T, c$	$t, c$
50	4369	18,3	0,366	7,7	0,154
100	9221	119,9	1,199	33,1	0,331
150	14073	416,0	2,773	85,2	0,568
200	18889	1082,6	5,413	177,6	0,888
250	23759	1800,2	7,201	274,5	1,098

Дополнительное повышение скорости работы модели потенциальных контейнеров достигается за счёт удаления ПК, которые не могут быть использованы для размещения ни одного из заданных объектов ни в одном из возможных вариантов их ориентации.

В табл. 2 сведены показатели скорости размещения наборов ОМ (рис. 2, а) при удалении неиспользуемых ПК и применении алгоритма 2 для обновления набора ПК. В последнем столбце таблицы содержится параметр  $K$ , показывающий, во сколько раз повышается скорость формирования упаковки при удалении неиспользуемых ПК.

**Таблица 2.**

**Скорость размещения наборов ОМ при удалении неиспользуемых ПК**

Число ОМ в задаче	Число ПК после размещения всех ОМ	Алгоритм 2		
		$T, c$	$t, c$	Повышение скорости размещения объектов ( $K$ )
50	3430	5,7	0,114	1,351
100	4444	21,8	0,218	1,518
150	5450	39,8	0,265	2,143
200	2559	57,8	0,289	3,073
250	3438	72,0	0,288	3,813

Полученные результаты тестирования демонстрируют значительное повышение скорости размещения объектов за счёт принципиального снижения числа обрабатываемых ПК. Важно отметить, что при удалении неиспользуемых ПК возрастает значение параметра скорости размещения объектов.

### Выводы

Разработанные алгоритмы оптимизации модели потенциальных контейнеров обеспечивают качественное повышение скорости геометрического конструирования упаковки объектов нерегулярной формы, представленных ортогональными многогранниками. В частности, применение алгоритма обновления набора ПК на основе операции пересечения, а также удаление неиспользуемых контейнеров для рассмотренной задачи размещения набора из 250 объектов позволило в 25 раз повысить скорость формирования упаковки.

Представленные в работе алгоритмы разработаны и программно реализованы с обобщением по размерности в авторском прикладном программном комплексе «Packer», предназначенном для решения задач раскроя и упаковки объектов сложной формы внутри контейнеров произвольной геометрии.

### Библиографический список

1. **Wäscher, G.** An improved typology of cutting and packing problems / G. Wäscher, H. Haußner, H. Schumann // *European Journal of Operational Research*. 2007. 183(3). P. 1109–1130.
2. **Mailloux, R.J.** Irregular polyomino-shaped subarrays for space-based active arrays / R. J. Mailloux, S. G. Santarelli, T. M. Roberts, D. Luu // *International Journal of Antennas and Propagation*. 2009. 956524. doi:10.1155/2009/956524.
3. **Bortfeldt, A.** Constraints in container loading - A state-of-the-art review / A. Bortfeldt, G. Wäscher // *European Journal of Operational Research*. 2013. 229(1). P. 1-20. doi:10.1016/j.ejor.2012.12.006.
4. **Byholm, T.** Effective packing of 3-dimensional voxel-based arbitrarily shaped particles / T. Byholm, M. Toivakka, J. Westerholm // *Powder Technology*. 2009. 196(2). P. 139-146.
5. **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.
6. **Arbuzov, M.O.** A progressive method of mounting machine parts on the shaft / M.O. Arbuzov, A.Ya. Nekrasov, A.N. Sobolev // *IOP Conf. Series: Materials Science and Engineering*. 2020. 709. 022066. doi:10.1088/1757-899X/709/2/022066.
7. **Chekanin, V.A.** Algorithms for management objects in orthogonal packing problems / V.A. Chekanin, A.V. Chekanin // *ARPN Journal of Engineering and Applied Sciences*. 2016. 11(13). P. 8436-8446.
8. **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.
9. **Bennell, J.** Tools of mathematical modeling of arbitrary object packing problems / J. Bennell, G. Scheithauer, Y. Stoyan, T. Romanova // *Annals of Operations Research*. 2020. 179(1). P. 343-368. doi:10.1007/s10479-008-0456-5.
10. **Stoyan, Y.** Optimized object packings using quasi-phi-functions / Y. Stoyan, T. Romanova, A. Pankratov, A. Chugay // In: G. Fasano, J. D. Pintér (Eds.), *Springer Optimization and Its Applications*. Springer, Cham. 2015. 105. P. 265–293. doi:10.1007/978-3-319-18899-7\_13.
11. **Chekanin, V.A.** Algorithm for the Placement of Orthogonal Polyhedrons for the Cutting and Packing Problems / V.A. Chekanin, A.V. Chekanin // In: Evgrafov A. (eds) *Advances in*

Mechanical Engineering. Lecture Notes in Mechanical Engineering. Springer, Cham. 2020. P. 41-48.

12. **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.
13. **Локтев, М.А.** Функциональный принцип обхода препятствий с применением метода функционально-воксельного моделирования / М.А. Локтев, А.В. Толоч // Вестник МГТУ «Станкин». 2016. № 1 (36). С. 75-80.
14. **Tolok, A. V.** Mathematical programming problems solving by functional voxel method / V. Tolok, N. B. Tolok // Automation and Remote Control. 2018. 79(9). P. 1703-1712. doi:10.1134/S0005117918090138.
15. **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.

## МЕТОДЫ АНАЛИЗА И КЛАССИФИКАЦИИ ИЗОБРАЖЕНИЙ В ПРОФОРИЕНТАЦИИ С УЧЕТОМ НЕЯВНЫХ СТРУКТУР И СМЕШАННОСТИ ТИПОВ

Ю.С. Тарасова

ORCID: 0000-0003-4770-9431 e-mail: tar06@list.ru

Нижегородский государственный архитектурно-строительный университет

В.В. Андреев

ORCID: 0000-0002-7557-352X e-mail: vyach.andreev@mail.ru

Нижегородский государственный технический университет им. Р.Е. Алексеева

Автоматизированная классификация изображений по выделенным типам в зависимости от наличия/отсутствия характерных искомых структур имеет огромное значение в повышении точности процесса профориентационного тестирования. Точность классификации способствует наиболее эффективной подготовке кадров, что является одним из существенных факторов снижения вероятности возникновения чрезвычайной или аварийной ситуаций, в частности – в инженерно-технической сфере, где человеческий фактор может стать причиной техногенной катастрофы. В статье рассматриваются автоматизированные профориентационные методики как элементы системы поддержки принятия решения, их достоинства и недостатки; нейросетевой и фильтрационный классификаторы как инструменты распознавания искомых структур (как явных, так и неявных), а также проблемы и перспективы процесса оптимизации и автоматизации методики профориентационного тестирования «Ассоциативное цветовое пространство» через программный комплекс ColourUniquePro.

**Ключевые слова:** поддержка принятия решения, классификация, нейросеть, фильтрация, профориентация

## METHODS OF IMAGE ANALYSIS AND CLASSIFICATION IN CAREER GUIDANCE, TAKING INTO ACCOUNT IMPLICIT STRUCTURES AND MIXED TYPES

I.S. Tarasova, V.V. Andreev

Nizhny Novgorod state university of architecture and civil engineering, Nizhny Novgorod

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

**Purpose:** The paper discusses the problem of identifying implicit structures and improving the accuracy of image classification in career guidance

**Design/methodology/approach:** For obtain the images, was used a quasi-spatial model like a part of the ColuorUniqueM software module. Part of the data obtained was used for training and validation during the development of a neural network classifier based on the architectures of three convolutional neural networks, another part of the data was classified using a filtering algorithm implemented by the «sliding window» method. The remaining part of the data was classified by using a «hybrid» classifier, that is, by jointly interpreting the classification results by a neural network classifier and a classifier based on a filtering algorithm.

**Findings:** The accuracy of machine classification has been improved, the process of collecting and analyzing data (images) has been automated, a series of experiments has been conducted for various types of tasks, such as optimization, classification, analysis of implicit structures, machine differentiation of subtypes, classification of mixed types.

**Research limitations/implications:** The research is limited by the diversity of the sample groups, and there is also a need to expand the classification criteria, in particular, a more detailed analysis of color characteristics.

**Originality/value:** The value of the research presented in the article consists in automating the methodology of career guidance testing, which allows generating unique images by working with a quasi-spatial model of the subject, which allows analyzing the individual characteristics of the individual's work with the color and structure of the quasi-space.

**Key words:** decision support, classification, neural network, filtering, career guidance



## Введение

В настоящее время необходимость актуализации профориентационных методик обусловлена автоматизацией и компьютеризацией практически всех сфер деятельности человека, в том числе – инженерно-технической.

Высокая вариативность возможностей и массивный объем разнообразной информации вызывает у современного абитуриента утомление и снижение мотивации, что оказывает влияние на скорость и качество принятия решений.

Существующие методы анализа предрасположенности абитуриентов (испытуемых) к той или иной сфере профессиональной деятельности являются широкодоступными, удобными и предоставляют испытуемому информацию по запросу определения наиболее подходящей сферы деятельности, на основе которой испытуемый принимает решение о выборе ВУЗа, после окончания обучения – о выборе профессии.

Существенными недостатками распространенных методик профориентационного тестирования являются устаревшие, неактуальные профессии в опросниках, сопоставление результатов тестирования по нескольким методикам может вызывать трудности интерпретации, в основном опросники ориентированы не на оценку его объективных качеств как потенциального специалиста, а на выявление его личных предпочтений [1].

При условии, что какая-либо профориентационная методика является автоматизированной, она может считаться системой поддержки принятия решений (СППР), так как в процессе тестирования испытуемые получают помощь в сложных условиях для полного и объективного анализа поступающих данных. При этом в процессе профориентационного тестирования и взаимодействия человека с интерфейсом программы, происходит извлечение и представление знаний, а также принятие решения, то есть – результат тестирования по результатам анализа и обработки полученных от пользователя данных[2].

В краткосрочной перспективе неактуализированный подход к профориентации приводит к прерыванию обучения будущим специалистом, недостаточному уровню подготовки кадров, низкой мотивации и включенности в будущую профессиональную деятельность, ошибкам, халатности, техногенным катастрофам, обусловленным человеческим фактором.

В данной статье рассматриваются методы анализа и классификации изображений в рамках методики тестирования «Ассоциативное цветовое пространство»©, оптимизированной и автоматизированной в виде программного комплекса для профориентации ColourUnique Pro. Методика предполагает отбор и визуализацию «чистых» типов, в также классификацию с учетом визуально неявных структур и смешанности типов с разграничением в процентном соотношении.

## Объекты исследования

Особенностью методики является уникальность изображения, получаемого в процессе работы пользователя (испытуемого) с квазипространственной моделью (рис. 1).

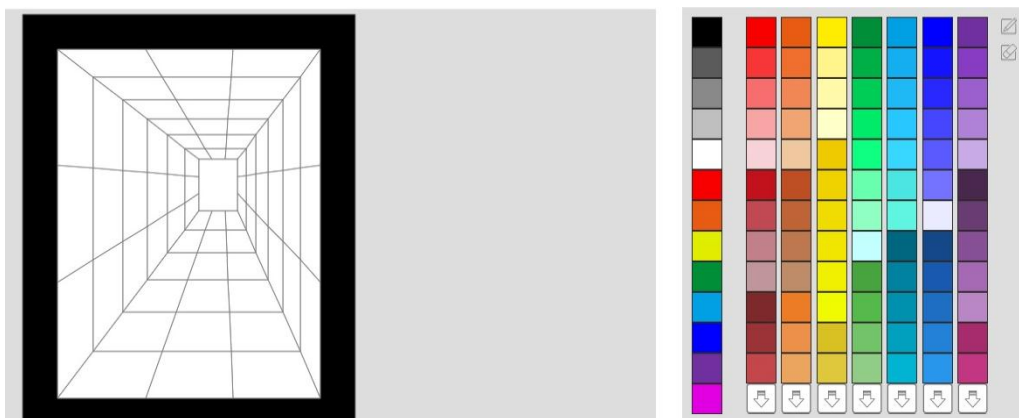


Рис.1. Слева – квазипространственная модель, справа – палитра для работы

Через интерфейс, в результате заполнения квазипространственной формы цветом, испытуемый генерирует уникальное изображение – цифровой «портрет», позволяющего характеризовать ИСД (индивидуальный стиль деятельности)[3], [4].

Так как пользователь (испытуемый) выполняет практическое задание (заполнение формы цветом), предполагающее получение индивидуального ответа, становится возможной оценка особенностей и подходов индивида к конкретной задаче, что и входит в понятие ИСД. По характеру работы с цветом и геометрией квазипространства, за время апробации методики, выявлено 6 характерных типов ИСД. Каждый тип отличается формированием особых элементов (см. рис. 2), которые являются искомыми в процессе автоматической машинной классификации.

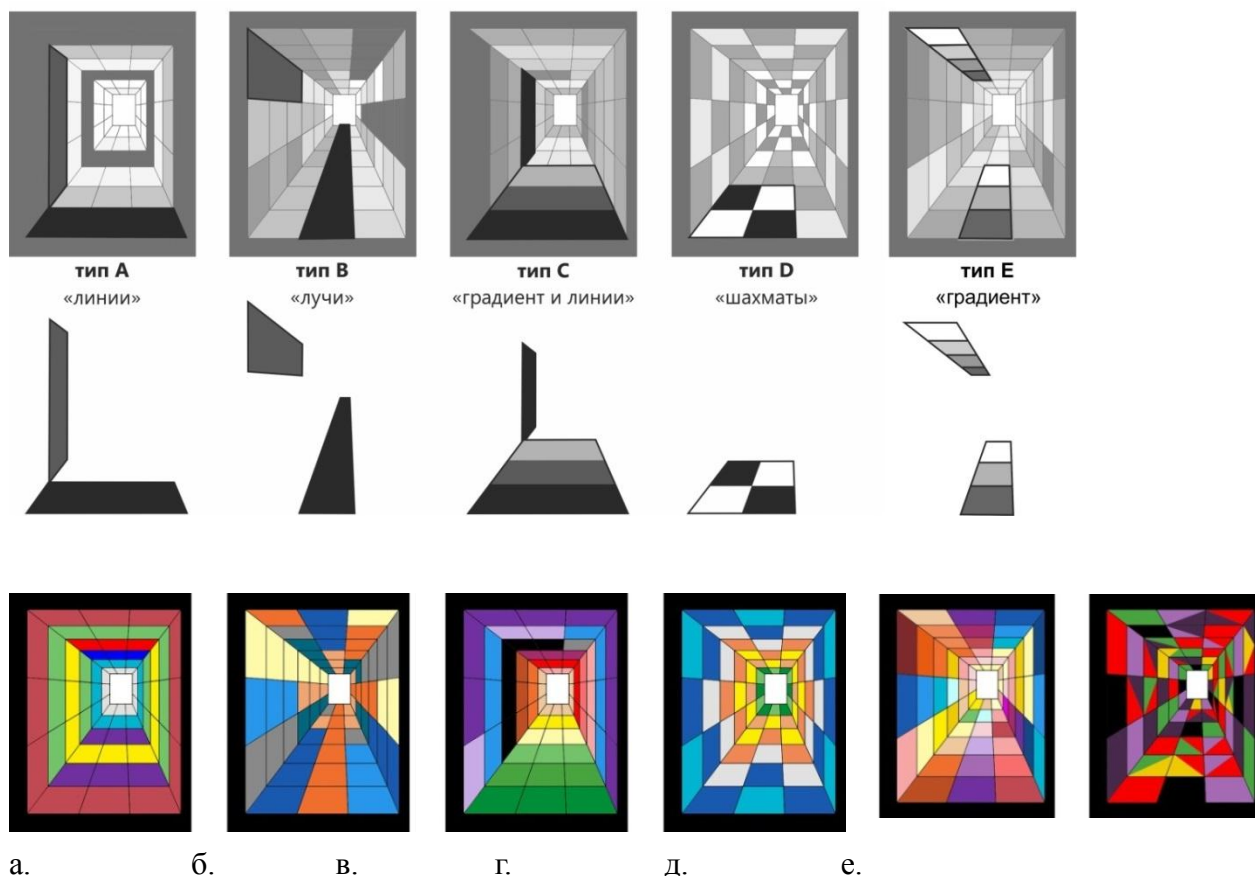


Рис. 2. – Выведенные типы индивидуалов: а. – тип А, б. – тип В, в. – тип С, г. – тип D, д. – тип E, е. – тип F

**Тип А** характеризуется преимущественным (от 50% и более, соответственно 100% - «чистый» тип) присутствием элементов «горизонтальные линии». Линии (от 3 ячеек в длину) могут составлять как цвета строго одинаковых координат, так и цвета разных координат, но относящихся к одному тону, например «зеленому», как на рис. 2, а.

**Тип В** характеризуется наличием (3 элемента на развертку и более) элементов «вертикальные линии» (от трех ячеек в высоту строго цвета одинаковых координат. Искомые элементы насчитывают минимум 3 ячейки в высоту, их может быть до 6, соответственно, область определения может достигать 6x1 (рис. 2, б).

**Тип С** характеризуется наличием (3 элемента на развертку и более) элементов «вертикальные линии» (от трех ячеек в высоту строго цвета одинаковых координат. Искомые элементы насчитывают минимум 3 ячейки в высоту, их может быть до 6, соответственно, область определения может достигать 6x1 (рис. 2, в).

**Тип D** характеризуется присутствием элементов «шахматы» или «подобие шахмат». Шахматные и шахматоподобные элементы выстраиваются за счет особого расположения темных и светлых ячеек в области определения 2x2. Причем шахматные элементы могут содержать любые цвета (рис. 2, г).

**Тип Е** главным образом характеризуется наличием большого количества оттенков выбираемых цветов, а именно – от 5 оттенков минимум для двух тонов. Как правило, их гораздо больше. Также некоторых представителей этого типа отличает наличие преимущественно «узких» градиентов, реже встречаются «широкие» (рис. 2, д).

**Тип F** является редким и предполагает работу с геометрией квазипространства через создание дополнительных линий (рис. 2, е). Это единственный класс изображений, определяющийся еще в процессе работы пользователя (испытуемого) с тестовой формой.

Пользователь (испытуемый) работает с формой, имеющей перспективные искажения, так как данный вид пространства помогает пользователю (испытуемому) понять суть задание и приступить к его выполнению.

Для машинного же анализа готовые изображения преобразуются в развертки, где каждая ячейка имеет вид квадрата, искажения отсутствуют (см. рис. 3).

### Методы анализа и классификации

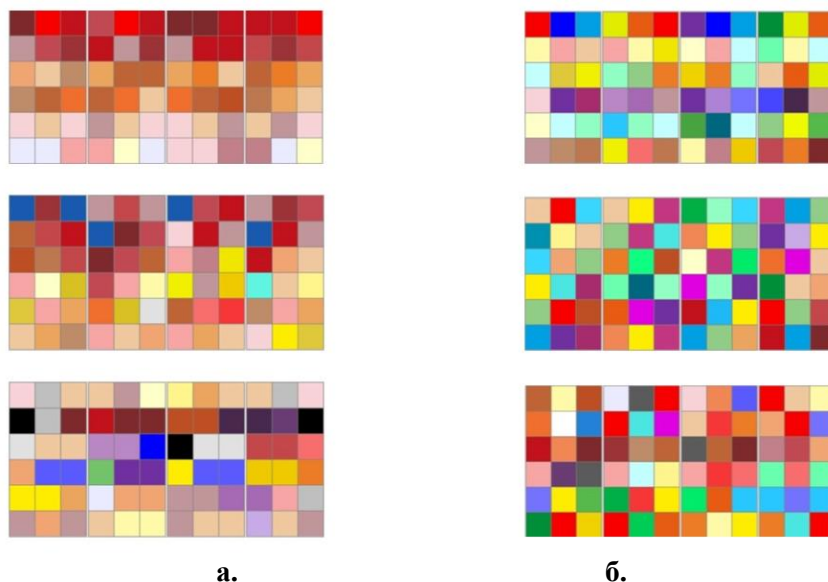
Развертки анализируются двумя классификаторами. Первый классификатор – нейросетевой, в основе которого находится архитектура сверточной сети InceptionV3 [4]. Второй классификатор – не нейросетевой, производящий алгоритмом фильтрации реализуемый на языке программирования Python за счет применения скользящего окна.

Нейросетевой классификатор показывает наибольшую эффективность в определении изображений классов А, С, и Е. Алгоритм фильтрации эффективен при обнаружении изображений классов В и D. Сопоставление результатов работы обоих классификаторов эффективно для автоматического разграничения подтипов а также дифференциации изображений, визуально «смешанных» по выраженности искомым структур.

Противоречия в классификации, в ряде случаев, позволяют выделить подтипы, либо указать для пользователя два наиболее возможных варианта, если дифференциация все же невозможна.

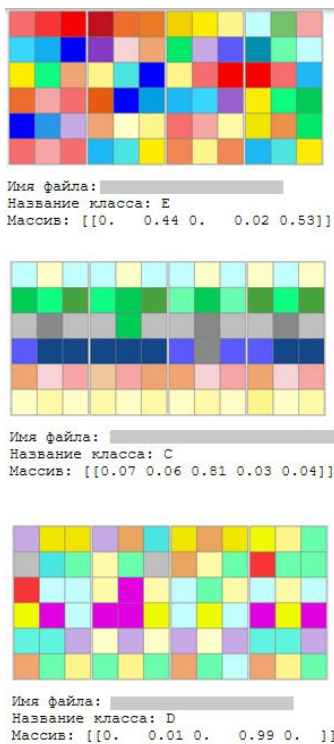
Рассмотрим сопоставление результатов классификации на примерах разграничения изображений, имеющих признаки как классов Е так и D, а также особенности идентификации типов А и В.

На рис. 3 изображены примеры визуально похожих изображений, относящиеся к разным искомым классам.

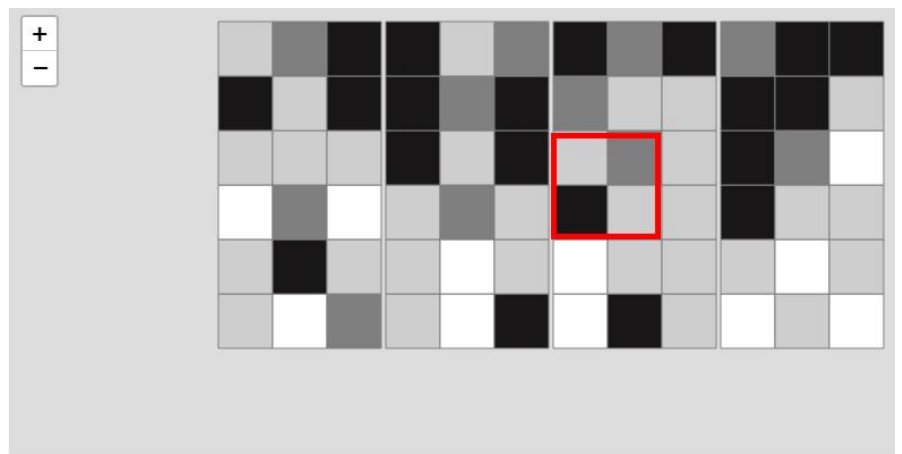


**Рис.3. а – изображения «неконтрастного» подтипа Е,  
б – изображения типа D с чертами типа Е**

Нейросеть классифицирует изображения на основе обучающей выборки, алгоритм фильтрации сканирует планигон [6] скользящим окном. Искомый элемент – «шахматоподобная» структура в области определения 2x2 (см. рис. 4, б)



а.



Цветовой диапазон: 103 127 181 Тёмный

Количество используемых цветов: 22

Всего квадратов: 55

Комбинации повторов 2-х состояний: 5 (9%)

Комбинации повторов 1 состояния: 9 (16%)

Комбинации без повторов: 4 (7%)

б.

**Рис.4. Результаты тестирования: а – нейросетевой классификатор, б – не нейросетевой классификатор**

Подсчет «шахматоподобных» элементов в процентном соотношении позволяет разделить подтипы следующим образом [7], [8]:



**Рис.5. Разделение подтипов D и E**

Искомymi элементами классов А и В являются горизонтальные и вертикальные линии соответственно, что, при стопроцентной выраженности, исключает присутствие «шахматоподобных» элементов класса D. Искомymi элементами классов С и Е считаются градиенты, и оба эти класса в целом визуально и структурно выглядят более сложными. Примеры разверток класса А, зачастую, не содержат других элементов, в то время как среди разверток классов С, D, или Е «чистые» типы встречаются редко, так как наличие градиентов в сочетании с линиями дает минимальное (до 10 %) присутствие «шахматоподобных» элементов, даже если пользователь (испытуемый) не формировал их намеренно.

Учитывая вышесказанное, можно утверждать, что при полном отсутствии «шахматоподобных» элементов (0 %) и соответствующих результатах после классификации через нейросеть, определяются «чистые» типы А и В:

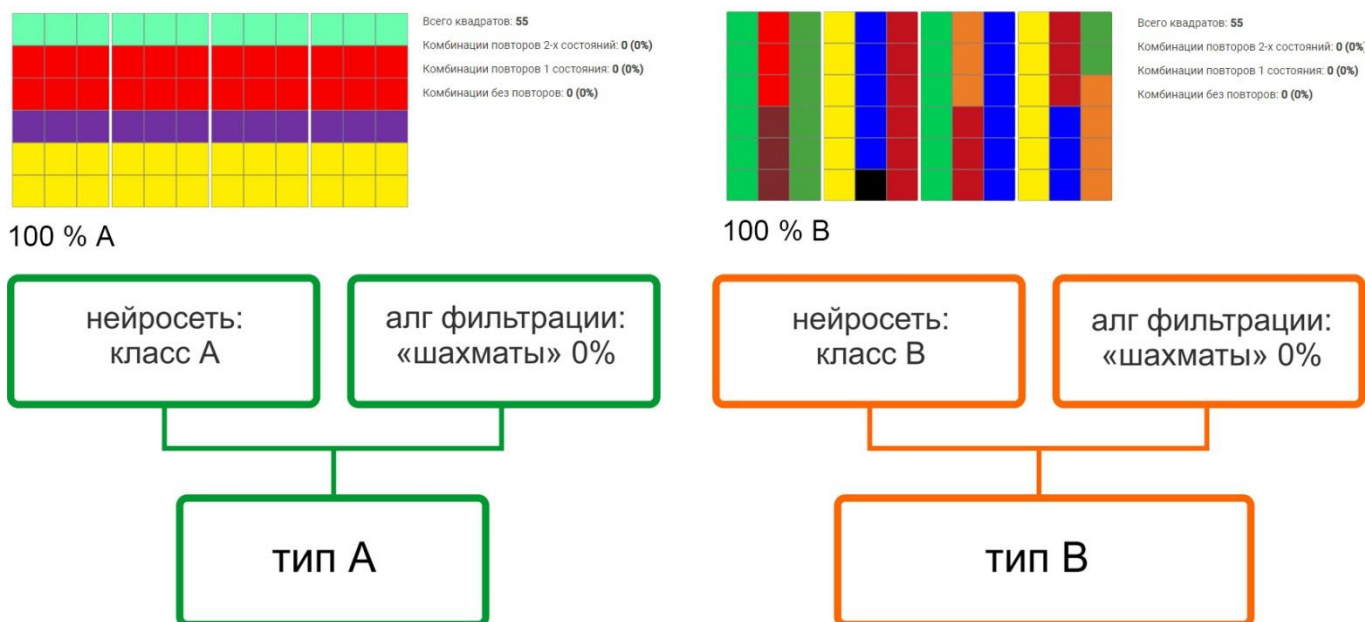


Рис.6. Разделение типов А и В

### Выводы

1. Разработаны и реализованы два классификатора: нейросетевой – на основе архитектуры сверточной нейронной сети, и не нейросетевой – с использованием алгоритма обработки изображения скользящим окном.
2. Для обоих классификаторов выявлены типы, определяемые с наибольшей точностью.
3. Сопоставлены и проанализированы результаты обработки изображений обоими классификаторами, что позволило повысить степень точности классификации а также дополнительно разграничить некоторые типы и подтипы.

### Библиографический список

1. Толстикова, А. А. Сравнительный анализ профориентационных методик / А. А. Толстикова. – Текст: непосредственный // Молодой ученый. – 2021. – № 17 (359). – С. 346-350. – URL: <https://moluch.ru/archive/359/80241/> (дата обращения: 08.03.2022).
2. Ларичев О.И., Петровский А.Б. Системы поддержки принятия решений: современное состояние и перспективы развития // Итоги науки и техники. М.: ВИНТИ, 1987. Т. 21, с. 131-164
3. Климов, Е.А. Индивидуальный стиль деятельности в зависимости от типологических свойств нервной системы / Е.А. Климов. – Казань: Изд-во Казан. ун-та, 1989. – 278 с.
4. Григорьева, М.В. Психология труда: конспект лекций / М.В. Григорьева – М.: Высшее образование, 2006 (Архангельск: ИПП Правда Севера). – 190 с.
5. С. Szegedy, W. Liu, Y. Jia, P. Sermanet, S. Reed, D. Anguelov, D. Erhan, V. Vanhoucke, A. Rabinovich, Going deeper with convolutions, 2014. URL: <https://arxiv.org/pdf/1409.4842.pdf>.
6. Утробин, В.А., Компьютерная обработка изображений. Информационные модели этапа понимания: учеб. пособие / В.А Утробин. – НГТУ. – Нижний Новгород, 2006. – 274 с.
7. Tarasova, I.S. Algorithms for automated differentiation of subtypes and improving the overall accuracy of image classification in career guidance / I.S. Tarasova, A.V. Chechin, V.V. Andreev// Computer Graphics and Vision. Proceedings of the 31th International Conference on Computer Graphics and Vision. Nizhny Novgorod, Russia, September 27-30, 2021. – P. 387- 398.
8. Тарасова, Ю.С. Автоматизированная методика профориентационного тестирования как система поддержки принятия решений [Текст] / Ю.С. Тарасова // Журнал «Системы управления и информационные технологии», №1(87), 2022. – С. 88-94.



## ПРИМЕНЕНИЕ МИВАРНОГО АЛГОРИТМА ДЛЯ РЕШЕНИЯ ЗАДАЧ РАСПРЕДЕЛЕНИЯ РЕСУРСОВ ПРОИЗВОДСТВЕННЫХ СИСТЕМ

О.В. Кривошеев

ORCID:0000-0002-9987-591X e-mail: OLvkrivosheev@rosatom.ru

Федеральное государственное унитарное предприятие «Российский федеральный ядерный центр – Всероссийский научно-исследовательский институт экспериментальной физики

Своевременное решение задач распределения ресурсов производственных систем в условиях неполноты данных играет важную роль как для действующих промышленных производств, так и для создания перспективного машиностроительного искусственного интеллекта и постепенного перехода к «безлюдным автономным производствам». Системный анализ описаний задач объемного планирования, объемно-календарного планирования и сменно-суточного планирования показал, что они могут быть формализованы и описаны в формате миварных сетей «объекты, правила», представляющих собой двудольные ориентированные графы. Возможность создания таких миварных моделей знаний обоснована тем, что подобные задачи были описаны и успешно решены для планирования действий киберфизических систем, автомобилей, различных роботов и робототехнических комплексов различного назначения и базирования. Для нахождения нужного варианта распределения ресурсов производственных систем, например, в виде общего плана предприятия предложено применять «Разуматор», который представляет собой новый механизм «без переборного» логического вывода на миварной базе знаний с линейной вычислительной сложностью. Для учета условий «неполноты данных» миварный алгоритм будет работать следующим образом. Новые ресурсы-объекты и операции-правила своевременно добавляются в миварную сеть в виде новых вершин и ребер двудольного графа с учетом анализа их свойств и ограничений. Недоступные ресурсы-объекты и операции-правила исключаются из миварной сети и затем, после проведения этих изменений в структуре миварной сети и учета всех новых ограничений и свойств, заново запускается логический поиск решения на основе «Разуматора» и определяется новый план распределения ресурсов производственных систем в виде маршрута логического вывода.

Работа выполнена в научной области системного анализа, управления и обработки информации в целях разработки и применения методов системного анализа сложных прикладных объектов исследования, обработки информации, целенаправленного воздействия человека на объекты исследования путем моделирования, оптимизации, совершенствования управления и принятия решений в целях повышения эффективности функционирования производственных систем. На основе миварных технологий логического искусственного интеллекта и методов оптимального планирования распределения ресурсов в сетевых канонических структурах разработан новый миварный алгоритм решения задач управления, принятия решений и обработки информации для своевременного распределения ресурсов производственных систем в условиях неполноты данных.

**Ключевые слова:** мивар, миварные сети, машиностроительный искусственный интеллект, экспертная система, MOGAN, MIPRA, КЭСМИ, Wi!Mi, Разуматор, Большие знания, распределение ресурсов производственных систем.

# MIVAR ALGORITHM FOR SOLVING PROBLEMS OF RESOURCE ALLOCATION OF PRODUCTION SYSTEMS IN CONDITIONS OF INCOMPLETE DATA

O.V. Krivosheev

FSUE RFNC-VNIIEF, Institute of Digital Technologies, Sarov, Russia

**Abstract.** Timely solution of the problems of resource allocation of production systems in conditions of incomplete data plays an important role both for existing industrial productions and for the creation of promising mechanical engineering artificial intelligence and the gradual transition to "deserted autonomous production". A systematic analysis of the descriptions of the tasks of volumetric planning, volumetric-calendar planning and shift-daily planning showed that they can be formalized and described in the format of mivar networks "objects, rules", which are bipartite oriented graphs. The possibility of creating such mivar knowledge models is justified by the fact that such tasks have been described and successfully solved for planning the actions of cyber-physical systems, cars, various robots and robotic complexes for various purposes and bases. To find the right option for allocating the resources of production systems, for example, in the form of a general enterprise plan, it is proposed to use a "Razumator", which is a new mechanism for "no-nonsense" logical inference on a mivar knowledge base with linear computational complexity. To take into account the conditions of "incomplete data", the mivar algorithm will work as follows. New resources-objects and operations-rules are added to the mivar network in a timely manner in the form of new vertices and edges of a bipartite graph, taking into account the analysis of their properties and constraints. Inaccessible resources-objects and operations-rules are excluded from the mivar network and then, after making these changes in the structure of the mivar network and taking into account all new restrictions and properties, the logical search for a solution based on the "Razumator" is restarted and a new plan for allocating resources of production systems is determined in the form of a logical output route.

The work was carried out in the scientific field of system analysis, management and information processing in order to develop and apply methods of system analysis of complex applied research objects, information processing, targeted human impact on research objects through modeling, optimization, improvement of management and decision-making in order to improve the efficiency of production systems. On the basis of mivar technologies of logical artificial intelligence and methods of optimal planning of resource allocation in network canonical structures, a new mivar algorithm for solving management problems, decision-making and information processing has been developed for the timely allocation of resources of production systems in conditions of incomplete data.

**Keywords:** mivar, mivar networks, mechanical engineering artificial intelligence, expert system, MOGAN, MIPRA, DSS, Wi!Mi, Razumator, decision support system, Big knowledge, resource allocation of production systems of enterprises.

## Введение

Проблема распределения ресурсов производственных систем играет важную роль при создании машиностроительного [1] искусственного интеллекта (МСИИ) [2]. В [3] приведен обзор зарубежных и отечественных научных работ, а также показан пример формализации широкого класса задач распределения ресурсов. К этим задачам относятся, например, оптимизационные задачи объемного планирования, объемно-календарного планирования и сменно-суточного планирования. Задачи оптимального планирования принято [3] рассматривать для производственных систем, функционирующих по такой схеме: изготовлению подлежит некоторая совокупность взаимозависимых работ, которые выполняются в дискретные моменты времени, называемые «такты планирования», при наличии в системе необходимых для их выполнения ресурсов. В свою очередь, ресурсы классифицируются по сроку годности на складываемые, не складываемые и частично-складываемые.

Важно отметить, что распределение ресурсов в этом случае моделируется сетевой канонической структурой - ориентированными взвешенными графами без петель и контуров, элементам которых поставлены в соответствие определенные характеристики. Как известно [3], к складываемым ресурсам относятся материалы, сырье, полуфабрикаты. Эти ресурсы могут быть использованы в любой такт функционирования системы, причем, если в некоторый такт они не использовались, то могут использоваться в любой последующий такт. К не складываемым ресурсам относятся трудовые ресурсы, фонд времени работы оборудования, транспортные средства. Эти ресурсы в случае, если они не были использованы в некотором такте планирования, в последующих тактах использоваться не могут. Частично-складываемые ресурсы (химикаты, скоропортящиеся продукты) могут быть использованы в заданные интервалы планирования [3].

Особенностью задач распределения ресурсов производственных систем (РПС) в условиях неполноты данных является то, что в определенные моменты времени некоторого ресурса может не оказаться в наличии, поэтому необходимо найти ему подходящую замену по определенным критериям. В случае когда замены нужного ресурса не существует приходится останавливать производство. Однако, существуют такие варианты, когда нужный ресурс можно заменить чем-то другим, например, некоторым ресурсом, у которого другие параметры, но в целом этот ресурс позволяет выполнить поставленную задачу и не останавливать производство. Так как распределение ресурсов в нашем случае моделируется ориентированными взвешенными графами без петель и контуров, то можно привести следующую аналогию с решением задач построения пути транспортных средств. Пусть граф состоит из вершин и ребер, а решение задачи распределения производственных ресурсов по аналогии представляет собой некоторый маршрут: последовательность вершин и ребер между ними. Тогда отсутствие некоторого ресурса можно моделировать либо отсутствием ребра – «ремонт дороги», «авария» и т.п., либо вершины – «закрытие перекрестка», «запрет въезда в город» и т.п. Конечно, это только упрощенная аналогия для улучшения наглядности процесса РПС. В этом случае нам необходимо найти альтернативный маршрут, который позволит связать начальную вершину с конечной вершиной. Как известно, такие задачи успешно решаются в теории графов или при построении логического вывода на миварных сетях [4].

Отметим, что важной особенностью рассматриваемой проблемы, является то, что расходование ресурсов может происходить с разными интенсивностями, от чего зависит и длительности выполнения работ. Тогда проблема оптимального планирования заключается в определении порядка выполнения заданной совокупности работ и стратегии распределения ресурсов между работами, при которых некоторые обобщенные показатели функционирования системы принимают экстремальные значения. Сложность подобных задач заключается как в большой размерности, так и в том, что необходимо полностью задать исходные параметры.



В настоящее время успешно применяют миварные технологии [4] логического искусственного интеллекта и созданные на их основе MOGAN [5] для различных классов задач. Миварные сети обуславливают переход к новому поколению экспертных систем и созданию интеллектуальных систем полного жизненного цикла изделий [1-2]. Миварные технологии [4] реализуют требования [6] обеспечения безопасности информации [7] в ключевых системах информационной инфраструктуры [8], а также могут встраиваться в сервисно-ориентированные архитектуры и многоагентные системы [9]. Миварные технологии являются сквозными цифровыми технологиями [10], которые применяются в медицине [11], например, для диагностики [12] сахарного диабета [13]; для создания интеллектуальных промышленных систем [14]; для понимания [15] компьютерами смысла текстов [16], речи [17] и образов [18]; для внедрения индивидуальных траекторий студентов вузов [19]; верификации аналитических и логических моделей [20]; а также для создания автономных транспортных систем [21], автомобилей [22], систем принятия решений («мозгов») для роботов [23] и групп робототехнических комплексов [24]; для подбора рекламных кампаний [25] на основе сравнения двух многомерных векторов в реальном времени [26] на основе миварных экспертных систем [27], а также для многих других применений, описанных в научной литературе. В [28] показаны основные преимущества миварного подхода по сравнению с подходами, основывающимися на онтологиях и когнитивных картах. Важно подчеркнуть, что для решения задач распределения ресурсов производственных систем в миварном подходе применяются классы и отношения [29], а также ограничения и прецеденты [30]. В 2021 году была опубликована новая методика создания миварных баз знаний [31] для решения практических задач в физическом мире, что также важно для перехода от задач логического вывода в дискретной математике (MOGAN) к созданию систем принятия решений для киберфизических систем, к которым можно отнести все производства, фабрики и заводы. Именно для автоматизации и цифровизации «физического производства» и создается «кибернетический» машиностроительный искусственный интеллект [1-2], чтобы в результате получить киберфизические «безлюдные производства». Столь широкий спектр решаемых задач позволяет перейти к исследованию возможности применения миварных сетей для оптимального планирования и решения задач распределения производственных ресурсов.

Обобщая полученную информацию, приходим к следующему выводу. Распределение ресурсов [3] моделируется сетевой канонической структурой - ориентированными взвешенными графами без петель и контуров, элементам которых поставлены в соответствие определенные характеристики. Проблема оптимального планирования [3] состоит в необходимости определить порядок выполнения заданной совокупности работ и найти такие стратегии распределения ресурсов между работами, при которых некоторые обобщенные показатели функционирования системы принимают экстремальные значения [3]. Следовательно, представляется целесообразным исследовать возможность применения миварных технологий (MOGAN) [4-31] для решения задач распределения ресурсов производственных систем, так как в этих разных подходах используются математические механизмы работы с ориентированными графами.

Таким образом, тема данной работы по созданию миварного алгоритма решения задач распределения ресурсов производственных систем в условиях неполноты данных актуальна и имеет важную практическую ценность.

### **Анализ задач распределения ресурсов и планирования с точки зрения применения миварных сетей**

Проблемы распределения ресурсов являются подклассом проблем теории расписаний. Решение данных задач является актуальным, так как они применяются в большом количестве различных сфер производственной деятельности. Например, типичными отраслями производства, где решают задачи распределения ресурсов, являются: управление производственной деятельностью; автоматизация процессов управления корпорацией;

разработка программного обеспечения; планирование научно-исследовательских и опытно-конструкторских работ; изготовление сложных изделий; строительство объектов; управление ресурсами многопроцессорного вычислительного комплекса при параллельных вычислениях и многие другие. Более подробный анализ широкого класса задач распределения ресурсов в сетевых канонических структурах приведен в работе [3]

Проведем краткое описание трех видов задач планирования для последующего анализа возможности создания миварного алгоритма решения этих задач. Прежде всего, для решения задачи объемного планирования требуется распределить общий план предприятия в объемных показателях - нормо-часах, рублях, условных тоннах. При этом, предполагается, что известными являются характеристики искомого плана, связанные с объемами работ по заказам, изделиям, тактам планирования. Формально задача объёмного планирования [3] ставится как задача определения таких объёмов работ, которые будут выполнены в каждом подразделении, по каждому заказу, каждому изделию, каждой детали в каждый такт планирования, для которых выполняются ограничения, связанные с заданными показателями искомого плана, и достигают экстремальные значения критерии, определяющие условия эффективного функционирования производственной системы. В формализме миварных сетей это можно описать соответствующим набором причинно-следственных правил с выделением в виде объектов - вершин двудольного графа: объемов работ; идентификаторов подразделений; идентификаторов заказов; идентификаторов изделий; идентификаторов деталей. Кроме того, отдельно описываются различные ограничения данной предметной области. Между всеми указанными выше объектами фиксируются правила перехода (миварные правила) – направленные ребра двудольного ориентированного графа.

Во втором случае, при решении задачи объемно-календарного планирования [3] рассматриваются производственные системы, в которых выполняется взаимозависимая совокупность работ. Каждая работа для своего выполнения должна пройти обработку на машинах согласно заданному для этой работы технологическому маршруту, который, в общем случае может быть произвольным. Для каждой работы определено время и затраты на ее выполнения на каждой машине согласно технологическому маршруту, соответствующему этой работе. Выполнение некоторых работ характеризуется не только потреблением ресурсов, необходимых для ее выполнения, но и производством ресурсов (полуфабрикатов), которые используются в дальнейшем процессе производства. Введение ресурсов, производимых в результате выполнения работ, позволяет описывать сетевую технологию изготовления изделий – для работы, начало выполнения которой не может начаться, пока не будут выполнены непосредственно ей предшествующие работы, эти работы являются «ресурсами», без наличия которых работа не может начать выполняться. При обычных для таких задач условиях (работа выполняется на машине без перерывов, на машине одновременно может выполняться не более одной работы), требуется построить такое расписание выполнения работ, которому соответствуют «наилучшие» оценки качества функционирования системы.

В формализме миварных сетей [4-25] это можно описать соответствующим набором причинно-следственных правил с выделением в виде объектов - вершин двудольного графа: машины, работы на машинах, технологические маршруты (и их составные части), интервалы времени, затраты на выполнение работы на каждой машине, производимые полуфабрикаты и др. Аналогично предыдущему случаю, отдельно описываются различные ограничения данной предметной области, а также между всеми объектами фиксируются правила перехода (миварные правила) – направленные ребра двудольного графа.

В третьем случае, в задачах сменно-суточного планирования [3] для каждой работы (операции) указана длительность ее изготовления и ресурс, который должен быть использован для ее выполнения (конкретный тип станка, агрегата, оборудования). Для каждого ресурса указано время его поступления в систему (для станков обычно это время начала планирования, или, например, время начала работы станка после ремонта). При этом,

условия, налагаемые на функционирование системы, разделяют на технологические, организационные и ресурсные. К технологическим условиям относятся условия взаимозависимости выполнения работ, условия, связанные с интенсивностями потребления работами ресурсов и условия, связанные с возможными длительностями выполнения работ. К организационным условиям относятся условия, связанные с моментами начала выполнения работ, условия, связанные с моментами окончания выполнения работ (директивные сроки). К ресурсным условиям относятся условия, связанные с расходом ресурсов при выполнении работ. Решением задачи сменно-суточного планирования является такое решение, для которого выполняются технологические, ресурсные и организационные условия, и достигается экстремального значения критерии, определяющие условия эффективного функционирования производственной системы [3]. Третий тип задач - сменно-суточного планирования в формализме миварных сетей также можно описать набором миварных правил с выделением в виде объектов - вершин двудольного графа: операции (работы), ее длительность, ресурсы для каждой операции, время поступления ресурса в систему, а также условия, которые в миварном формализме называют «ограничения». Ограничения будут тоже трех видов: технологические, организационные и ресурсные. Аналогично предыдущим случаям, между всеми объектами фиксируются правила перехода (миварные правила) – направленные ребра двудольного графа.

Анализ кратких описаний задач объемного планирования, объемно-календарного планирования и сменно-суточного планирования показал, что все они могут быть формализованы и описаны в требуемом формате миварных сетей: объекты и правила. Обоснованием для утверждения о возможности создания таких миварных моделей будет являться тот факт, что с некоторой степенью аналогии подобные задачи были описаны и успешно решены для планирования действий робототехнических комплексов различного назначения в научных работах [4-25]. Важно отметить, что после создания адекватной модели предметной области, для нахождения различных вариантов решений по распределению общего плана предприятия будет применяться механизм логического вывода на миварной базе знаний [4-25]. Также специально подчеркнем, что вопросы поиска оптимального решения исследуются в других научных работах и выходят за рамки данного исследования. Целью данной работы является обоснование возможности создания миварного алгоритма решения задач распределения ресурсов производственных систем в условиях неполноты данных.

#### **Обоснование возможности создания миварного алгоритма решения задач распределения ресурсов производственных систем в условиях неполноты данных**

При планировании распределения ресурсов производственных систем принято опираться на полное описание всех «цепочек» для необходимого построения плана. Однако возникают такие ситуации, когда некоторые ресурсы или операции перестают выпускаться или выполняться, а вместо них начинают использоваться некоторые другие ресурсы и операции. Например, могут измениться материалы или свойства некоторых ресурсов или у них появятся другие характеристики. Регулярно происходят изменения станков, агрегатов и другого оборудования. Важно, что из «номенклатуры изделий» некоторое изделие перестанет выпускаться, а для выполнения плана необходимо будет искать замену этому изделию. Аналогично может изменяться и состав ресурсов – оборудования предприятия. Возможно, придется искать замену не только одному изделию/станку, но и целой группе таких изделий/станков с условием выполнения требуемых характеристик.

Важно зафиксировать тот факт, что надо будет в явном виде формализовать и выделить важные характеристики изделия/оборудования, которые точно должны быть сохранены или улучшены, а также определить перечень остальных характеристик, которые не критичны и могут быть изменены достаточно сильно.

С формальной точки зрения, в таком случае возникают условия «неполноты данных», когда по существующему ранее плану работ (технологической карте, схеме) нельзя создать

требуемое изделие и необходимо вносить изменения в планирование распределения ресурсов и оборудования производственных систем.

В качестве решения проблемы планирования (распределения ресурсов) в условиях «неполноты данных» предлагается разработать миварный алгоритм решения этих задач на основе использования технологий создания и применения миварных экспертных систем с логическим выводом на основе «Разуматора» [1-2, 4-25]. В таком случае предварительно необходимо создать соответствующую модель знаний «ресурсов производственных систем».

Выше было обосновано, что миварные модели знаний можно создавать для задач объемного планирования, объемно-календарного планирования и сменно-суточного планирования, так как все они могут быть формализованы в требуемом формате миварных сетей: объекты и правила «Если – То» - двудольного ориентированного графа. В таких создаваемых моделях знаний будут в явном виде фиксироваться (запоминаться) правила применения различных ресурсов со всеми их разновидностями и параметрами. Кроме того, в такой модели знаний будут своевременно фиксироваться различные новые виды ресурсов (и операций с ними).

Как было показано, для нахождения плана распределения ресурсов предприятия могут использоваться миварные сети, в которых собрана вся необходимая информация и создана требуемая база знаний для предметной области. Для решения задач планирования в условиях «неполноты данных» целесообразно применять миварные сети и базы знаний.

#### **Миварный алгоритм решения задач распределения ресурсов производственных систем в условиях неполноты данных**

Миварный алгоритм решения задач распределения ресурсов производственных систем (РРПС) в условиях неполноты данных будет основан на следующих принципах работы. Выше было предложено построить аналогию с решением задач построения пути транспортных средств, т.к. распределение ресурсов в нашем случае моделируется ориентированными взвешенными графами. Итак, граф состоит из вершин и ребер, а решение задачи распределения производственных ресурсов по аналогии представляет собой некоторый маршрут: последовательность вершин и ребер между ними [32-34]. Приведены примеры: отсутствие некоторого ресурса можно моделировать либо отсутствием ребра – «ремонт дороги», «авария» и т.п., либо вершины – «закрытие перекрестка», «запрет въезда в город» и т.п. Конечно, это только упрощенная аналогия для улучшения наглядности процесса РРПС. В этом случае нам необходимо найти альтернативный маршрут, который позволит связать начальную вершину с конечной вершиной.

Необходимо добавить, что миварная сеть – это двудольный ориентированный граф, а значит все вершины разбиты на два множества: «объектов» и «правил» [33]. Для задачи РРПС под «объектами» будем понимать материальные ресурсы: детали, изделия и т.п., а под «правилами» будем понимать процессы, например, как переход от заготовки к детали, которую изготавливают на определенном станке. Так же правила будут отображать процесс сборки из различных деталей некоторого изделия, а затем и сборку более сложного изделия из нескольких более простых изделий. Миварный подход [32] позволяет проводить системное описание любой предметной области [34] и вводить некие правила перехода с одного уровня системного описания на другой.

Например, в результате нескольких сложных правил был собран двигатель автомобиля, а также аналогично были подготовлены кузов и другие сложные комплектующие изделия этого системного уровня для поставки на конвейер. Тогда процесс сбора самого автомобиля из комплектующих предыдущего системного уровня: двигатель, кузов и т.п., будет представлять собой правило перехода на следующий системный уровень, где из одних изделий собирается уже конечное изделие – автомобиль.

У каждого объекта и правила миварной сети есть свой соответствующий набор свойств, которые могут выполнять и функции ограничений. Например, если на конвейер по сбору автомобилей не поступили какие-либо комплектующие, то и правило «сбор

автомобиля» не будет запускаться, а система выдаст сообщение о нехватке входных комплектующих.

Фактически, миварный алгоритм решения задач распределения ресурсов производственных систем (РРПС) в условиях неполноты данных представляет собой логический вывод на миварной базе знаний (рис. 1), т.е. на заранее созданной миварной сети предметной области РРПС. Сам по себе логический вывод выполняется автоматически путем использования «Разуматора» - специально созданного программного обеспечения в виде миварного логического «движка» [27], который подробно описан в учебных пособиях [32-34] с различными примерами применения. А для использования миварного Разуматора необходимо создать миварную сеть для каждой предметной области решения задач распределения ресурсов производственных систем.

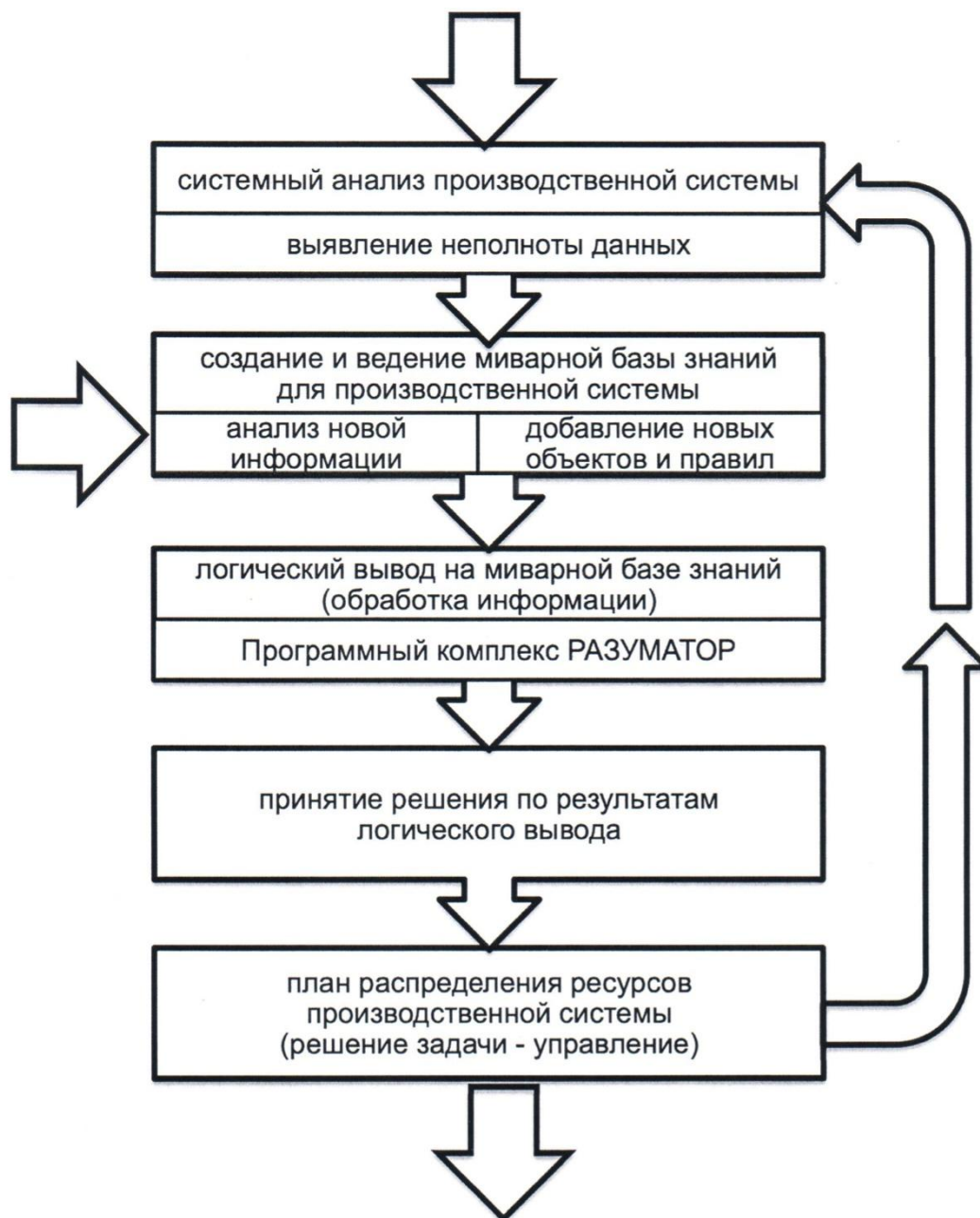


Рис. 1. Схема работы миварного алгоритма решения задач РРПС.

Таким образом, миварный алгоритм решения задач распределения ресурсов производственных систем в условиях неполноты данных, показанный на рис.1, можно представить следующим образом. На первом этапе вход системы поступает информация о

производственной системе и проводится ее системный анализ. На втором этапе на основании результатов системного анализа создается миварная база знаний по конкретной производственной системе. Кроме того, здесь же на втором этапе производится сбор и анализ новой информации, которая может понадобится в случае выявления неполноты данных. Пока вся необходимая информация есть система работает в штатном режиме на 1 этапе. В случае выявления неполноты данных, примеры которой приведены выше, переходим на второй этап в подпункт удаления «неисправных» или «отсутствующих» объектов и правил, а также к добавлению в миварную сеть новых объектов и правил.

Далее переходим к этапу 3: загружаем имеющуюся миварную базу знаний в программный комплекс Разуматор, задаем начальные и требуемые объекты, а затем запускаем Разуматор.

На следующем 4 этапе Разуматор выдает полученный алгоритм логического вывода или сообщает о нехватке входных данных, а система переходит к принятию решения по результатам работы Разуматора. Система является автоматизированной и человек-эксперт анализирует полученные результаты и принимает решение, которое и передает на 5 этап.

На 5 этапе логический вывод преобразуется в план распределения производственной системы и передается на исполнение для проведения процессов управления производственной системой. Вместе с этим, результаты работы передаются и на самый первый этап для системного анализа и принятия решений на следующие периоды принятия решения. Таким образом, происходит дополнение данных для решения задач распределения ресурсов производственных систем [3]. А сами условия выявления неполноты добавляются новой информацией. Алгоритм завершает свою работу и опять переходит в режим ожидания выявления новых фактов неполноты данных. В этом же фоновом режиме возможно постоянное выполнение 2 этапа по анализу новой информации и внешних условиях производственной системы и возможных способах устранения неполноты данных.

Теоретически решение задач РРПС можно описать следующим образом (рис. 1). Если некоторый ресурс становится недоступным (по разным причинам), то в миварной базе такой ресурс исключается из миварной сети и в дальнейшем не используется. Существует возможность на исходном этапе решения задач изменить начальные условия или структуру миварной сети, путем исключения или добавления новых ребер и вершин. В случае «неполноты данных» недоступные ресурсы исключаются из миварной сети, а новые ресурсы и операции своевременно добавляются в миварную сеть в виде новых вершин и ребер двудольного графа. Далее, после проведения изменений в структуре миварной сети, вновь запускается поиск решения на основе «Разуматора» и находится новый план распределения ресурсов. Некоторые ресурсы/операции являются критическими и их исключение потребует значительного изменения миварной базы знаний, чем и будут заниматься люди-эксперты с привлечением аналитиков и когнитологов – специалистов по созданию миварных баз знаний.

Приведем пример по предметной области «сборка автомобиля». Допустим, для заводской комплектации автомобиля используются некоторые шины марки X1, которые производитель шин перестал выпускать в некоторый момент времени. Теперь вместо шин марки X1 стали выпускать шины марки X2. Однако, существуют и другие производители шин, которые выпускают марки шин А, В, С и т.д. Понятно, что пример специально подобран так, чтобы решение было найти достаточно просто: в миварной сети, например, на складе или при закупке шин появляется сообщение, что шины марки X1 больше не будут поставляться. И вершина с маркой X1 удаляется из миварной сети. Вместе с тем, запускается специальная процедура поиска замены шин марки X1 на какую-нибудь другую марку шин, но которая должна полностью удовлетворять всем требованиям, например, по размерам, надежности и до стоимости таких шин. Такая процедура оформляется в миварной сети в виде некоторого правила поиска замены одного промежуточного изделия (или детали) на другое изделие/деталь. Если удастся найти такую замену, которую контролирует человек, то в миварной сети появляется новая вершина с заменой марки шин (на X2, или А1, или В1 и т.д.). Отметим, что в миварной сети можно будет добавить несколько новых вершин с

марками шин, которые прошли процедуры проверки соответствия. Например: А1, А2, В1, В3 и т.д.). Необходимо отметить, что любая такая замена вершины будет приводить к пересчету, например, себестоимости изделия – автомобиля, причем как в большую, так и в меньшую сторону.

Можно привести пример замены вершины типа «правило». Например, некоторую операцию выполняют на некотором станке, которая и описывается в виде «правила» перехода от заготовки к детали. Пусть в некоторый момент времени, такой станок переводят на ремонт или он совсем ломается и требуется его замена. В миварной сети правило, связанное с этим станком (или несколько правил), сразу переводится в режим «недоступно». Здесь возможны разные варианты, но основными будут:

- 1) ремонт станка и возврат правила в «доступ»;
- 2) удаление правила, если станок списывают совсем;
- 3) замена правила на другой станок, если такой находят или покупают новый станок и правило возвращается в режим «доступно».

Возможны и другие варианты. Например, самый простой вариант, если таких станков несколько и можно один станок заменить на другой – равнозначный. Это возможно в случае, если один станок переходит в режим ремонта, а другой наоборот, возвращают из ремонта. Более сложный вариант, когда аналогичный станок можно найти в другом помещении, районе или даже городе: понятно, что в таком случае увеличиваются времена доставки заготовок к станку и потом перевозка нужной детали от станка к месту использования этой детали. Учитывая развитие промышленности, возможны варианты, когда некоторые типы изделий вообще перестают выпускать самостоятельно, а переходят к их закупке/поставке с других предприятий. Это возможно при создании корпораций, покупке новых предприятий и т.п. В таком случае миварная сеть описания задачи РРПС должна пройти адаптацию для перехода к новым условиям.

Важно отметить, что в любом случае, миварная сеть позволяет решать задачи поиска новых алгоритмов и маршрутов в условиях неполноты данных по аналогии с тем, как такие задачи решал бы человек на основании логических рассуждений и своих знаний [1-2, 4-31]. Напомним, что базовый формат миварных правил – это причинно-следственные или каузальные правила формата «Если..., То ...». Хотя для миварных сетей предложены различные варианты развития самого формализма представления знаний [28-30].

Возможны различные ситуации и случаи, но обосновано, что использование миварных сетей позволит решить некоторое количество новых случаев и позволит создать комбинированные алгоритмы решения задач распределения ресурсов производственных систем в условиях неполноты данных. Комбинированность алгоритмов – это совместное использование традиционно применяемых алгоритмов для теории расписаний с добавлением к ним миварных технологий логического искусственного интеллекта, а прежде всего, миварных экспертных систем на основе логического «Разуматора» - автоматического построения алгоритмов решений в миварных сетях, представляющих собой двудольные ориентированные графы.

### **Выводы**

Решение задач распределения производственных ресурсов в условиях неполноты данных играет значительную роль для создания машиностроительного искусственного интеллекта. Работа выполнена в научной области системного анализа, управления и обработки информации в целях разработки и применения методов системного анализа сложных прикладных объектов исследования, обработки информации, целенаправленного воздействия человека на объекты исследования путем моделирования, оптимизации, совершенствования управления и принятия решений в целях повышения эффективности функционирования производственных систем. На основе миварных технологий логического искусственного интеллекта и методов оптимального планирования распределения ресурсов в сетевых канонических структурах разработан новый миварный алгоритм решения задач

управления, принятия решений и обработки информации для своевременного распределения ресурсов производственных систем в условиях неполноты данных.

Анализ описаний задач объемного планирования, объемно-календарного планирования и сменно-суточного планирования показал, что они могут быть формализованы и описаны в формате миварных сетей «объекты, правила», представляющих собой двудольные ориентированные графы. Возможность создания таких миварных моделей обоснована тем, что подобные задачи были описаны и успешно решены для планирования действий робототехнических комплексов различного назначения.

Предложен новый миварный алгоритм решения задач распределения ресурсов производственных систем в условиях неполноты данных. Для нахождения различных вариантов распределения общего плана предприятия будет применяться «Разуматор» - механизм логического вывода на миварной базе знаний. Для условий «неполноты данных» миварный алгоритм будет работать следующим образом: новые ресурсы и операции своевременно добавляются в миварную сеть в виде новых вершин и ребер двудольного графа с учетом анализа их свойств и ограничений, а недоступные ресурсы исключаются из миварной сети и затем, после проведения изменений в структуре миварной сети, заново запускается поиск решения на основе «Разуматора» и находится новый план распределения ресурсов производственных систем.

#### **Библиографический список**

1. **Варламов, О.О., Кривошеев, О.В., Трищенко, А.В., Осипов, В.Г., Лялин, Е.С.** О новом научном направлении исследований "машиностроительный искусственный интеллект" для систем полного жизненного цикла изделий // В сборнике: труды международного научно-технического конгресса "интеллектуальные системы и информационные технологии - 2021" ("ИС & ИТ-2021", "IS&IT'21"). 2021. С. 84-91.

2. **Варламов, О.О., Кривошеев, О.В., Трищенко, А.В., Осипов, В.Г., Лялин, Е.С.** О научном направлении «Машиностроительный искусственный интеллект» для систем полного жизненного цикла изделий // Фундаментальные основы физики, химии и механики наукоемких технологических систем формообразования и сборки изделий : сборник трудов научного симпозиума технологов- машиностроителей – ЭБС ДГТУ. – С. 227-235.

3. **Прилуцкий, М.Х., Власов, В.С., Кривошеев, О.В.** Задачи оптимального планирования как задачи распределения ресурсов в сетевых канонических структурах // Информационные технологии. 2017. Т. 23. № 9. С. 650-657.

4. **Варламов, О.О.** Эволюционные базы данных и знаний для адаптивного синтеза интеллектуальных систем. Миварное информационное пространство. - М.: Радио и связь, 2002. - 288 с.

5. **Варламов, О.О.** Обзор 18 миварных экспертных систем, созданных на основе MOGAN // Известия Кабардино-Балкарского научного центра РАН. 2021. № 3 (101). С. 5-20.

6. **Варламов, О.О.** О системном подходе к созданию модели компьютерных угроз и ее роли в обеспечении безопасности информации в ключевых системах информационной инфраструктуры // Известия ТРТУ. 2006. № 7 (62). С. 216-223.

7. **Кузьменко, Г.Н., Амарян, М.Р., Адамова, Л.Е. и др.** Роль интеллектуальных систем информационной безопасности для Рунета // Искусственный интеллект. 2005. №4. С. 757-762.

8. **Блохина, С.В., Адамова, Л.Е. и др.** Разработка учебных программ с элементами искусственного интеллекта для обучения в области информационной безопасности и защиты персональных данных // Искусственный интеллект. 2009. № 3. С. 328-335.

9. **Варламов О.О.** Анализ взаимосвязей GRID И САС ИВК, SOA И миварного подхода // Искусственный интеллект. 2005. № 4. С. 4-11.

10. **Varlamov, O.O., Chuvikov, D.A., Adamova, L.E. et al.** Logical, Philosophical And Ethical Aspects Of AI In Medicine // International Journal of Machine Learning and Computing. 2019. Т. 9. № 6. С. 868-873.



11. **Колесниченко, О.Ю., Мартынов, А.В., Пулит, В.В. и др.** Современный передовой уровень искусственного интеллекта для умной медицины // Ремедиум. Журнал о российском рынке лекарств и медицинской техники. 2019. № 4. С. 36-43.
12. **Ким, Х., Чувиков, Д.А., Аладин, Д.В. и др.** Создание базы знаний для миварной экспертной системы диагностики сахарного диабета // Медицинская техника. 2020. № 6 (324). С. 38-41.
13. **Белоусов, Е.А., Попов, И.А., Евдокимов, А.А. и др.** Рекомендательная система диагностики сахарного диабета на основе механизма миварного вывода // Естественные и технические науки. 2021. № 7 (158). С. 169-174.
14. **Санду, Р.А.** Миварный подход к созданию интеллектуальных систем и искусственного интеллекта. Результаты 25 лет развития и ближайшие перспективы. - Москва, 2010.
15. **Варламов, О.О., Адамова, Л.Е., Елисеев, Д.В. и др.** О миварном подходе к моделированию процессов понимания компьютерами смысла текстов, речи и образов. Новые возможности расширения границ автоматизации умственной деятельности человека // Автоматизация и управление в технических системах. 2013. № 2 (4). С. 38-51.
16. **Адамова, Л.Е., Сурикова, О.В., Булатова, И.Г. и др.** Применение миварной экспертной системы для оценки сложности текстов // Известия Кабардино-Балкарского научного центра РАН. 2021. № 2 (100). С. 11-29.
17. **Варламов, О.О.** Миварные технологии как некоторые направления искусственного интеллекта // Проблемы искусственного интеллекта. 2015. № 1 (1). С. 23-37.
18. **Варламов, О.О.** Автоматизация умственной деятельности людей через логический искусственный интеллект как фундаментальный механизм развития или гибели человечества // Проблемы искусственного интеллекта. 2017. № 3 (6). С. 23-31.
19. **Адамова, Л.Е.** Применение миварных технологий для внедрения в инженерном и экономическом образовании индивидуальных траекторий студентов // Известия Кабардино-Балкарского научного центра РАН. 2020. № 1 (93). С. 18-34.
20. **Назаров, К.В.** Разработка методики создания верифицируемых моделей для миварных экспертных систем // T-Comm: Телекоммуникации и транспорт. 2017. Т. 11. № 4. С. 64-71.
21. **Aladin, D.V., et al.** Logic-based artificial intelligence in systems for monitoring the enforcing traffic regulations // В сборнике: IOP Conference Series: Materials Science and Engineering. Moscow, 2019. С. 012025.
22. **Varlamov, O.O., Chuvikov, D.A., Aladin, D.V., et al.** Logical artificial intelligence mivar technologies for autonomous road vehicles // В сборнике: IOP Conference Series: Materials Science and Engineering. Moscow, 2019. С. 012015.
23. **Varlamov, O.O.** "Brains" for Robots: Application of the Mivar Expert Systems for Implementation of Autonomous Intelligent Robots // Big Data Research. 2021. Vol. 25, 100241. <https://doi.org/10.1016/j.bdr.2021.100241>.
24. **Варламов, О.О.** О создании на основе миварных систем принятия решений "РОБО!РАЗУМ" групп автономных комбайнов и тракторов для сельского хозяйства // Проблемы искусственного интеллекта. 2019. № 2 (13). С. 49-62.
25. **Семенов, А.А.** Исследование способов подбора рекламных кампаний на основе сравнения многомерных векторов // Проблемы искусственного интеллекта. 2020. № 1 (16). С. 89-104.
26. **Семенов, А.А.** Разработка метода сравнения двух многомерных векторов в реальном времени на основе миварных экспертных систем // Проблемы искусственного интеллекта. 2020. № 2 (17). С. 94-109.
27. **Varlamov, O.O.** Wi!Mi expert system shell as the novel tool for building knowledge-based systems with linear computational complexity // International Review of Automatic Control. 2018. Vol. 11. № 6. P. 314-325. <https://doi.org/10.15866/ireaco.v11i6.15855>.

28. **Чибирова, М.О.** Сравнительный анализ миварного подхода с подходами, основывающимися на онтологиях и когнитивных картах // Радиопромышленность. 2015. №3. С. 55-66.
29. **Чибирова, М.О.** Структурное развитие миварного подхода: классы и отношения // Радиопромышленность. 2015. № 3. С. 44-54.
30. **Чибирова М.О.** Необходимость добавления ограничений и прецедентов для развития миварного информационного пространства // Радиопромышленность. 2015. № 3. С. 66-78.
31. **Chuvikov, D.A., Aladin, D.V., Adamova, L.E.et al.** A new method for creating Mivar know ledgebasesintabular-matrixform forground intelligent vehicle controlsystems. [Journalof Physics: Conference Series](https://doi.org/10.1088/1742-6596/2061/1/012123), 2021, 2061(1) 012123. <https://doi.org/10.1088/1742-6596/2061/1/012123>.
32. **Варламов, О.О.** 18 примеров миварных экспертных систем. М.: ИНФРА-М, 2021. 630 с.
33. **Варламов, О.О.** Основы создания миварных экспертных систем. Учебное пособие / М.: ИНФРА-М, 2021. — 267 с.
34. **Варламов, О.О.** Миварные базы данных и правил. Учебное пособие / М.: ИНФРА-М, 2021. — 351 с.

## ПРОГНОЗИРОВАНИЕ С ПОМОЩЬЮ МОДЕЛИ ЗНАНИЙ ОСД

А.С. Макаров

ORCID: 0000-0002-4468-7161 e-mail: alexmakjob@yandex.ru

Нижегородский государственный технический университет им. Р.Е. Алексеева

В работе предложен алгоритм оценки контента с помощью модели знаний ОСД. Оценка позволяет найти более эффективные факторы влияющие на систему, даёт возможность улучшения или удаления ненужных элементов системы. При наличии коллекции проектных планов оценка может служить фактором ранжирования, которая позволит выбрать лучшую альтернативу для реализации. В исследовании производилась преобразование контента в сеть ОСД для последующего расчёта эмоционального коэффициента. Для получения дополнительных данных использовался метод экспертной оценки. Проведён количественный и качественный анализ существующих алгоритмов с целью определения их достоинств и недостатков. В статье рассматривается исследование прогнозирования успешности контента с помощью знаний ОСД.

**Ключевые слова:** анализ текстов, искусственный интеллект, модели знаний, анализ медиаконтента, экспертная система.

ANALYSIS OF THE QUALITY OF R-PEAK DETECTORS  
IN ECG SIGNAL IN THE REAL-TIME MODE

A.S. Makarov

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

**Purpose:** The article is devoted to the study of an ability of research with the help of the OPA model.

**Design / methodology / approach:** The article deals with the content characteristics before its creation. The author proposes transformation of the content into an OPA network and then to conduct an expert survey and calculate the emotional coefficient. This metrics can help a company to choose the best project plan and save the company's expenses.

**Findings:** An ability of the usage of emotional coefficient is also present in the article.

**Research implications:** The emotional coefficient can't be a reliable assessment of the potential profitability of content because it does not have an option of predicting external systems factors.

**Originality:** Such emotional coefficient lets estimate OPA network, textual or visual content at the development stage. Evaluation lets you find more effective factors affecting the system, makes it possible to identify or remove unnecessary elements from the system itself. A collection of project plans, the emotional coefficient score can serve as a ranking factor to select the best alternative to implement. Collection of project plans and the evaluation of the emotional coefficient can serve as a ranking factor that will allow you to choose the best alternative for implementation. Forecasting using OPA lets you get an initial assessment at the product development stage which significantly saves the company's expenses. Analysts no longer need to spend their time on developing evaluation criteria. To add to this, the probability of choosing the wrong method or lack of the necessary criteria disappear. The OPA network can serve as a language for describing or developing which can let see the evaluation of the proposed solutions.

**Key words:** text analysis, artificial intelligence, knowledge models, media content analysis, expert system.

## Введение

Технологический прогресс не стоит на месте. Конкуренция на рынке вынуждает компании внедрять новые оценочные метрики для анализа потенциала визуального или текстового контента до его выпуска. Как правило, популярность контента связана с количеством эмоций [11] получаемых потребителем. Из этого возникает потребность в оценке контента для повышения вероятности контентного успеха.

Известны методы оценки применимых для анализа контента: экспертный опрос, сетевой анализ, теория полезности [9]. Однако каждая теория имеет свои достоинства и недостатки.

В настоящее время для оценки контента используется экспертный опрос [2]. Организатор опроса собирает группу независимых экспертов. Они оценивают продукт по определённым критериям. Перечень критериев [9] определяет организатор опроса. После чего производится расчёт и подведение результатов.

Данный тип опросов имеет ряд недостатков:

1. для проведения опроса требуется контент;
2. сложность определения критериев;
3. отсутствие точного языка описания контента.

В данной статье исследуется возможность использования модели знаний ОСД для оценки выпускаемого контента. Прогнозирование с помощью ОСД относится к эвристическому методу прогнозирования.

## Этапы прогнозирования

Этапы прогнозирования с помощью модели ОСД:

1. выбор цели;
2. классификация сущностей для модели ОСД;
3. построение сети ОСД;
4. получение экспертных оценок частоты появления объектов;
5. получение экспертных оценок частоты взаимодействий связей объектов;
6. расчёт эмоционального коэффициента.

## Прогнозирование с помощью модели ОСД

Исследуемый контент должен представлять систему взаимодействующих элементов. Он может быть представлен, как видео поток данных или текстовый материал. Для преобразования контента в модель знаний ОСД следует использовать (см. табл. 1).

Для получения метрики оценки контента, потребуется построить сеть ОСД.

ОСД аббревиатура означает: О-объект, С-свойство, Д-действие. ОСД – модель знаний, описывает взаимодействие объектов в системе.

Данную модель знаний можно представить в ориентированном и неориентированном мультиграфе. Вершина графа – объект, рёбра или дуги-действие. Свойство - это дополнительный элемент системы, который может быть связан с множеством объектов (рис. 1).

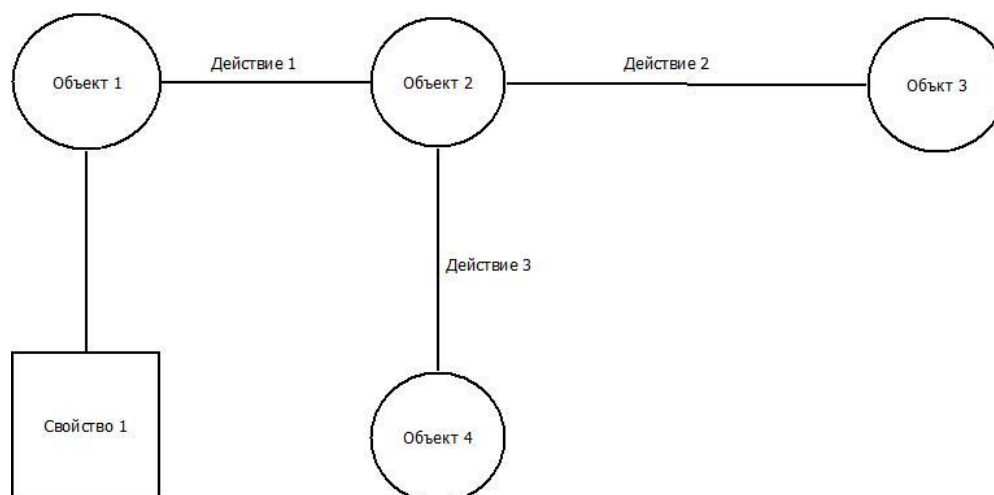


Рис. 1. Модель ОСД

Таблица 1.

## Преобразование информации в модель знаний ОСД

Текстовый контент	Визуальный контент	Роль в модели ОСД
Имя существительное	Визуальный объект	объект
Имя числительное	Исключён	исключён
Имя прилагательное	Свойство объекта, которое можно определить визуально.	свойство
Местоимение	Визуальный объект	метаданные
Глагол	Взаимодействие двух объектов	действие
Наречие	Свойство объекта, которое можно определить визуально.	свойство
Предикатив	Свойство объекта, которое можно определить визуально.	свойство
Причастие	Свойство объекта, которое можно определить визуально.	свойство
Деепричастие	Свойство объекта, которое можно определить визуально.	свойство
Предлог	исключён	исключён
Союз («и», «или», также символ «,»)  при условии формулы: [модельный тип] [,] [модельный тип]	N визуальный объект	логические метаданные
Частица	Свойство объекта, которое можно определить визуально.	свойство
Модальные слова	исключён	исключён
Междометия	исключён	исключён

Зададим треугольную матрицу  $\Theta$ , в которой  $\lambda_{ij} = 0$  при  $i > j$  и  $\lambda_{ij} = 1$  при  $i < j$ . Данная матрица позволит убрать из расчётов дубликации связей и усреднений частот объектов.

Для расчётов необходимо получить частоту появления объектов  $\nu_{объект}$  из оценок экспертов  $\rho$ . Эксперту по очереди предъявляется объект из модели ОСД. Задача эксперта: оценить частоту появления объекта в данной предметной области. Он выбирает вариант ответа от 0 до 10 (см. табл. 2). Эксперту может быть представлен диапазон времени, который поможет ему с поиском упоминаний данного объекта в коллекции, существующих контентов.

Таблица 2.

## Преобразование оценок экспертов

Объективное восприятие	$\rho_{эксперта}$
Очень часто	10
	9
	8
	7
	6
	5
	4
	3
	2
Не видел никогда	1

$$v_{\text{объект}} = \varrho/100, \quad (1)$$

По данным формируется матрица частот объектов вида:

$$\begin{pmatrix} v_1 \\ v_2 \\ \dots \\ v_n \end{pmatrix},$$

(2)

где  $n$  – количество объектов.

$$Z = \begin{pmatrix} K_{11} & K_{12} & K_{13} & \dots & K_{1j} \\ K_{21} & K_{22} & K_{23} & \dots & K_{2j} \\ K_{31} & K_{32} & K_{33} & \dots & K_{3j} \\ \dots & \dots & \dots & \dots & \dots \\ K_{i1} & K_{i2} & K_{i3} & \dots & K_{ij} \end{pmatrix}, \quad (3)$$

где  $k_{ij}$  – количество рёбер мультиграфа.

$$K = Z \text{ mult } \Theta, \quad (4)$$

Необходимо получить матрицу смежности по усреднённым частотам появления объектов с общей связью,  $\alpha_{ij}$  – средняя частота соседних объектов.

$$\alpha_{ij} = \frac{(v_{\text{текущего объекта}} + v_{\text{соседнего объекта}})}{2},$$

если связь отсутствует, то в матрице указывается  $\alpha_{ij}$  равен 0.

$$O = \begin{pmatrix} \alpha_{11} & \alpha_{12} & \alpha_{13} & \dots & \alpha_{1j} \\ \alpha_{21} & \alpha_{22} & \alpha_{23} & \dots & \alpha_{2j} \\ \alpha_{31} & \alpha_{32} & \alpha_{33} & \dots & \alpha_{3j} \\ \dots & \dots & \dots & \dots & \dots \\ \alpha_{i1} & \alpha_{i2} & \alpha_{i3} & \dots & \alpha_{ij} \end{pmatrix}, \quad (5)$$

$$A = O \text{ mult } \Theta, \quad (6)$$

Для расчётов необходимо получить частоту появления двух связанных объектов и действие  $\omega$  из оценок экспертов  $\varrho$ . Эксперту по очереди предъявляется предложение в виде из модели ОСД. Задача эксперта: оценить частоту появления связи в данной предметной области. Он выбирает вариант ответа от 0 до 10 (см. табл. 2). Эксперту может быть представлен диапазон времени, который поможет ему с поиском упоминаний данной связи в коллекции, существующих контентов.

Расчёт частоты связей объектов  $\omega_{ij}$  из оценок эксперта  $\varrho_{\text{эксперта}}$  по формуле (7).

$$\omega_{ij} = \varrho/100, \quad (7)$$

Если два объекта имеют более одной связи, то  $\omega_{ij}$  производится расчёт среднего арифметического связей:

$$\omega_{ij} = \frac{\sum \omega_c}{k_{ij}} \quad (8)$$

где  $\omega_c$  – частота связи объекта с общими объектами.

По данным формируется матрица частот связей вида:

$$E = \begin{pmatrix} \omega_{11} & \omega_{12} & \omega_{13} & \dots & \omega_{1j} \\ \omega_{21} & \omega_{22} & \omega_{23} & \dots & \omega_{2j} \\ \omega_{31} & \omega_{32} & \omega_{33} & \dots & \omega_{3j} \\ \dots & \dots & \dots & \dots & \dots \\ \omega_{i1} & \omega_{i2} & \omega_{i3} & \dots & \omega_{ij} \end{pmatrix}, \quad (9)$$

$$M = E \text{ mult } \Theta, \quad (10)$$

Расчёт эмоциональной матрицы:

$$\Phi = (A \text{ mult } K) - (M \text{ mult } K), \quad (11)$$

Расчёт эмоционального коэффициента:

$$\Omega = \frac{\sum \Phi}{s}, \quad (12)$$

где  $s$  – количество связей.

$$s = \sum K, \quad (13)$$

$\Omega$  – эмоциональный коэффициент, где  $\Omega \in [-0.9, 0.9]$

## Результаты исследования существующих проектов

В результате исследования было произведен анализ мобильных приложений. В генеральную выборку попали коммерчески успешные и провальные проекты. В результате исследования была разработана программа автоматического расчета эмоционального коэффициента (рис. 2).

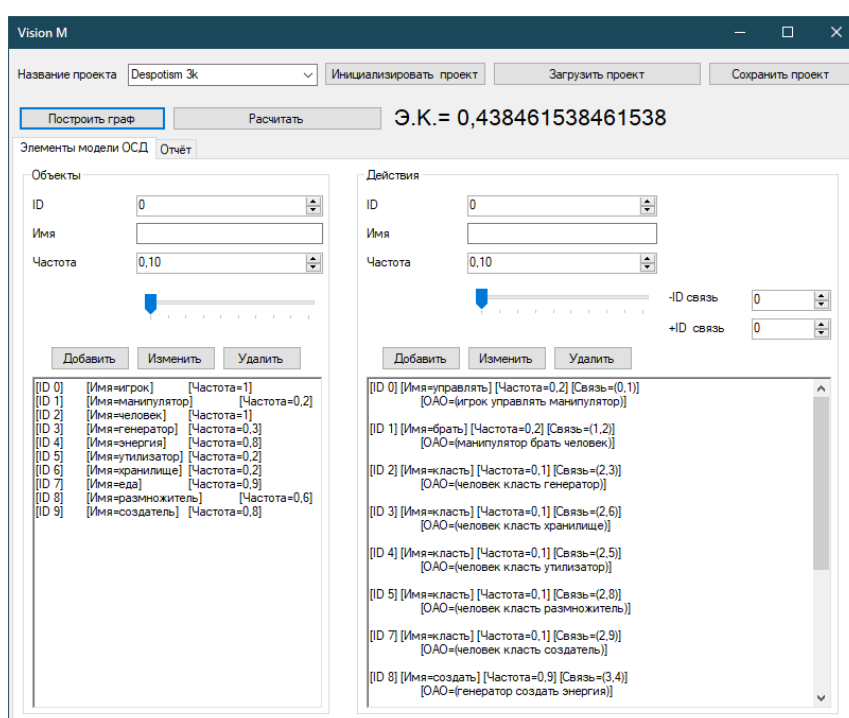


Рис. 2. Расчетная программа эмоционального коэффициента

Результаты исследования эмоциональных коэффициентов проектов были занесены в таблицу (см. табл. 3).

Из ранее полученных результатов можно выделить контентные группы. Контенты, обладающие одинаковым эмоциональным коэффициентом, имеют схожие показатели выручки. Однако данный вид сравнения нельзя считать правдоподобным, так в системе явно присутствуют прочие факторы, а именно влияние медиаресурсов. Прогнозирование с помощью модели ОСД предсказывает следующий шаг итерации динамической системы, но плохо обосновывает существующие данные.

Результат расчётов эмоциональных коэффициентов проектов и выручки компании

Название проекта	Эмоциональный коэффициент	Выручка (руб.)
Alien Blackout	0,18	14500000
Despotism 3k	0,43	14900000
Earn to Die	0,48	15000000
Five Nights at Freddy	0,35	239000000
Goat Simulator	0,33	379000000
Horror evil	0,025	7500
Peace Death	0,5	15000000
Peace Death 2	0,06	245000
Время пушки	0,19	35000
Тук-тук-тук	0,275	990000

### Заключение

Данный эмоциональный коэффициент позволяет оценить ОСД сеть, текстовый или визуальный контент на этапе его проектирования. Оценка позволяет найти более эффективные факторы влияющие на систему, даёт возможность улучшения или удаления ненужных элементов системы. При наличии коллекции проектных планов оценка эмоционального коэффициента может служить фактором ранжирования, которая позволит выбрать лучшую альтернативу для реализации. Прогнозирование с помощью ОСД позволяет получить первичную оценку, на этапе проектирования продукта, что существенным образом экономит траты компании. Аналитикам ненужно тратить время на формирования критериев оценки, а также исчезает вероятность ошибки выбора неправильного или нехватка нужного критерия. Сеть ОСД может служить языком описания и проектирования, которая позволит в динамике видеть оценку предлагаемых решений.

### 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.
9. **Ларичев, О.И.** Теория и методы принятия решений, а также Хроника событий в Волшебных Странах: Учебник. - М.: Логос, 2000. - 296 с : ил.
10. **Kenna, Ralph, Mac Carron, Padraig.** Maths Meets Myths: Network Investigations of Ancient Narratives // Journal of Physics: Conference Series. — 2016. — Vol. 681. — P. 012002.
11. **Симонов, П. В.** Эмоциональный мозг. — М.: Наука, 1981. — 215 с.



## РЕКОМЕНДАТЕЛЬНАЯ СИСТЕМА С ПОМОЩЬЮ ОСД ОНТОЛОГИИ

А.С. Макаров

ORCID: 0000-0002-4468-7161 e-mail: alexmakjob@yandex.ru

Нижегородский государственный технический университет им. Р.Е. Алексеева

Интеллектуальные системы помогают человеку в большом количестве сфер человеческой жизни. Однако человеку иногда сложно найти в информационной системе нужную информацию или функциональность. Решением такой проблемы занимается подсистема рекомендаций. Данная система прогнозирует релевантность контента на базе поведенческих данных человека. В статье рассматривается аналитическое сравнение существующих рекомендательных систем с рекомендательной системой, основанной на ОСД онтологии. Также раскрываются недостатки существующих систем. ОСД аббревиатура означает: О – объект, С – свойство, Д – действие. ОСД – модель знаний, описывает взаимодействие объектов в системе. Данная онтология позволяет описывать, как текстовые данные, так и визуальные, что даёт возможность гибкого использования. ОСД онтология может организовывать подход коллаборативной фильтрации item based, делая прогноз на базе опыта, а также реализовать подходы content-based и knowledge-based. Особое внимание уделено рекомендательной системе: описаны алгоритмы и принципы работы. Так как ОСД онтология удовлетворяет классическому определению системы, такой взгляд на проблему рекомендательных систем будет интересен специалистам в области искусственного интеллекта.

**Ключевые слова:** рекомендательная система, искусственный интеллект, модели знаний.

## RECOMMENDER SYSTEM USING OPA ONTOLOGY

A.S. Makarov

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

**Purpose:** The article is devoted to an analytical comparison of recommender people and an overview of the recommender system based on the OPA ontology.

**Design / methodology / approach:** The article provides an analytical overview of existing systems. For each type of recommender system, the disadvantages and advantages of such approaches are formulated. For a recommender system based on the ontology of the OPA, algorithms are described and an analytical review is made.

**Findings:** The advantages of a recommender system based on the OPA ontology are determined.

**Research implications:** The advantages of a recommender system based on the OPA ontology are determined. Development of work algorithms for a recommender system based on OPA ontology.

**Originality:** The originality of the work is due to the description and formation of the idea of using the OPA ontology for recommender systems. As a result of the analytical review, it was found that such a system has a number of advantages over existing approaches, one of which is the ease of implementation in other systems. Particular technical advantages are: ease of use of OPA ontology, accounting for behavioral factors, ease of calculation, personalized recommendations, a small amount of data to work with, the system takes into account the shown content recommendations.

**Key words:** recommender system, artificial intelligence, knowledge models.

## Введение

Рекомендательные системы — программы, которые пытаются предсказать, какие объекты будут интересны пользователю, основываясь на данных пользователя или внешних источниках [1].

Есть две основные стратегии создания рекомендательных систем — фильтрация на основе содержания и коллаборативная фильтрация.

При фильтрации на основе содержания создаются профили пользователей и объектов. Профили пользователей могут включать демографическую информацию или ответы на определённый набор вопросов. Профили объектов могут включать названия жанров, имена актёров, имена исполнителей и другую атрибутивную информацию в зависимости от типа объекта.

При коллаборативной фильтрации используется информация о поведении пользователей в прошлом — например, информация о покупках или оценках. В этом случае не имеет значение, с какими типами объектов ведётся работа, но при этом могут учитываться неявные характеристики, которые сложно было бы учесть при создании профиля.

В процессе работы рекомендательные системы собирают данные о пользователях.

Примеры оценочных данных:

1. оценки объекта по дифференцированной шкале;
2. ранжировки группы объектов от наилучшего к наихудшему;
3. сравнение двух объектов и выбор лучшего;
4. ранжированные по приоритету.

Примеры косвенных данных:

1. запись событий;
2. ведение записей о поведении пользователя онлайн;
3. отслеживание содержимого компьютера пользователя.

Рекомендательные системы сравнивают и анализируют данные от разных пользователей и вычисляют список рекомендаций для конкретного пользователя. Рекомендательные системы — удобная альтернатива поисковым алгоритмам, так как позволяют обнаружить объекты, которые не могут быть найдены последними.

## Типы рекомендательных систем

Существует 4 типа рекомендательных систем:

1. Коллаборативная фильтрация (collaborative filtering) [14];
2. Основанные на контенте (content-based) [1];
3. Основанные на знаниях (knowledge-based) [13];
4. Гибридные (hybrid) [1].

## Обзор типов рекомендательных систем

Преимущество коллаборативной фильтрации:

1. учёт оценок пользователей;
2. подход обладает теоретически высокой точностью;
3. автоматическое создание рекомендаций.

Недостатки коллаборативной фильтрации:

1. сложность обучения;
2. ошибки обучения нейронной сети;
3. отсутствие возможности гибкой настройки;
4. высокая квалификация специалистов.

Преимущество основанных на контенте:

1. прогнозирование на базе эмпирического опыта;
2. сбор данных о пользователе.

Недостатки основанных на контенте:

1. сильная зависимость от предметной области;
2. ограниченность полезности рекомендаций;
3. снижение точности.

Преимущество основанных на знаниях:

1. рекомендации основаны на знаниях о предметной области;
2. высокую точность;
3. учёт индивидуальных свойств;
4. анализирует взаимосвязи между объектами;
5. пользовательские пожелания;
6. демографические особенности.

Недостатки основанных на знаниях:

1. сложность создания;
2. высокая квалификация специалистов.

### **Формирование сети ОСД для рекомендательной системы**

Для создания рекомендательной системы необходимо создать сеть ОСД. ОСД сеть – является онтологией. ОСД аббревиатура означает: О – объект, С – свойство, Д – действие. ОСД – модель знаний, описывает взаимодействие объектов в системе. Данная онтология позволяет описывать, как текстовые данные, так и визуальные, что даёт возможность гибкого использования.

Инженеру необходимо определить элементы системы и возможные взаимодействия, которые позволят создать рекомендации для пользователя. После определения элементов системы и взаимодействий необходимо сформировать ОСД сеть, используя таблицу преобразования в модель знаний ОСД(см. табл. 1).

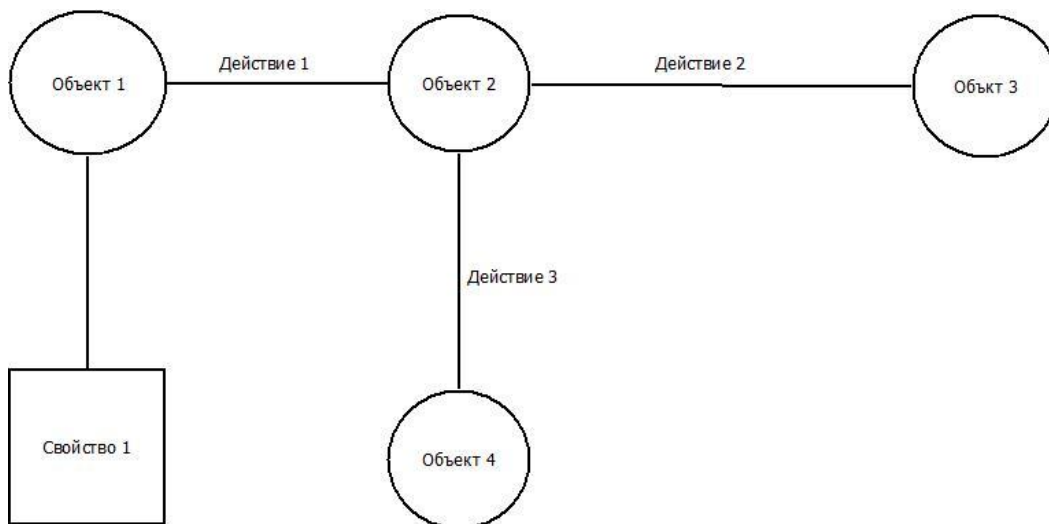
В качестве объекта в модели ОСД может выступать каталог, веб - страница, вкладка программы, элемент программы. В качестве действия в модели ОСД может выступать любое действие: переход, показать, выдвинуть.

**Таблица 1.**

#### **Преобразование информации в модель знаний ОСД**

<b>Визуальный контент</b>	<b>Роль в модели ОСД</b>
Визуальный объект	объект
Свойство объекта, которое можно определить визуально.	свойство
Визуальный объект	метаданные
Взаимодействие двух объектов	действие
N визуальный объект	логические метаданные

Данную модель знаний можно представить в ориентированном и неориентированном мультиграфе. Вершина графа – объект, рёбра или дуги - действие. Свойство - это дополнительный элемент системы, который может быть связан с множеством объектов (рис. 1).



**Рис. 1. Модель ОСД**

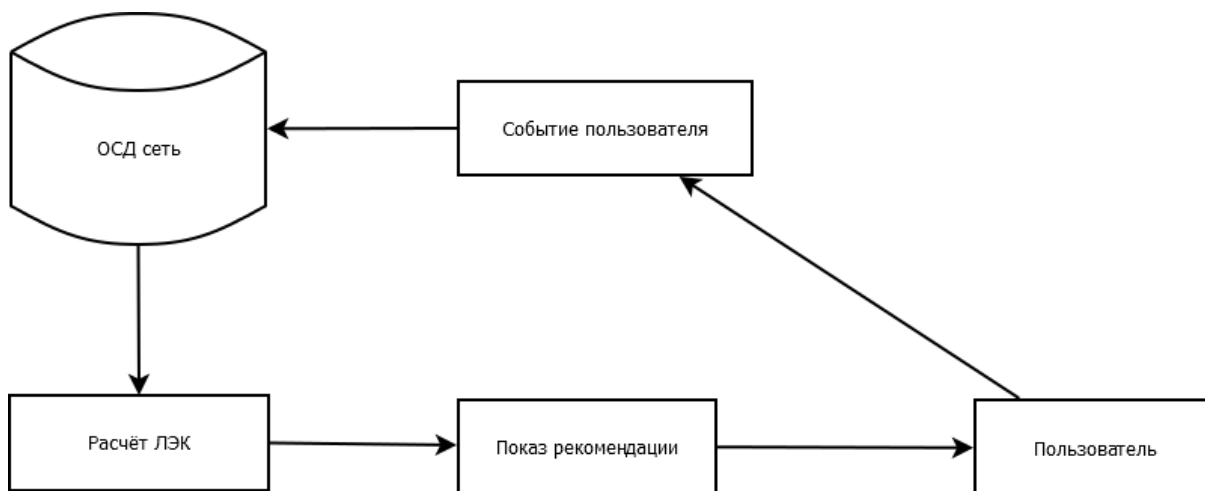
### Принцип рекомендательной системы ОСД

Рекомендательной системы с использованием ОСД онтологии представляет, из себя интегративную, динамическую систему с временной памятью. Работа такой системы разделена на несколько этапов.

Этапы работы:

1. сбор данных;
2. расчёт ЛЭК, формирование списка приоритета рекомендаций;
3. показ рекомендации.

ЛЭК – локальный эмоциональный коэффициент, который служит мерой оценки триплета объект, действия, объект.



**Рис. 2. Схема работы рекомендационной системы ОСД**

Инженер определяет время работы сбора данных от пользователя, также устанавливает пороговое значение для показа рекомендации.

По окончании времени все коэффициенты ОСД сети обнуляются и данные собираются заново.

Пользователь взаимодействует с информационной системой, она собирает данные и отправляет их в сеть ОСД. Если пользователь увидел объект, который описан в сети ОСД, то показатель  $v_{объект}$  увеличивается на 1. Где  $v_{объект}$  количество показов объекта. Если пользователь

совершает взаимодействие, которое описано в сети ОСД, то  $\omega_{\text{действия}}$  увеличивается на 1. Где  $\omega_{\text{действия}}$  количество повторяемых действий. После каждого события, происходит расчёт ЛЭК для действия «рекомендовать», если произошло увеличение число показов на объекте по формуле 1.

$$\text{ЛЭК} = \alpha - \omega_{\text{действия}}, \quad (1)$$

Формула усреднения показов объектов:

$$\alpha = \frac{(v_{\text{текущего объекта}} + v_{\text{соседнего объекта}})}{2}, \quad (2)$$

Если объект имеет несколько связей «рекомендовать», то производится ранжирование по параметру ЛЭК от большего к меньшему и выбирается связь «рекомендовать» с большим значением ЛЭК.



**Рис. 3. Пример ОСД онтологии для рекомендательной системы**

### **Преимущества рекомендательной системы ОСД**

Преимущества рекомендательная система ОСД:

1. количество показов объектов можно увеличивать из внешних источников;
2. система учитывает показанные рекомендации контента;
3. простота расчётов для ранжирования;
4. учёт поведенческих факторов;
5. возможность управления рекомендациями;
6. широкие возможности по внедрению;
7. возможность интеграции дополнительных ОСД сетей;
8. гибкость учёта метрик;
9. персонализированные рекомендации;
10. простота внедрения в другие системы
11. прогнозирование действий пользователей.

Недостатки рекомендательная система ОСД:

1. наличие в команде инженера знаний;
2. отсутствие возможности показа рекомендации при отсутствии достаточных данных.

### **Вывод**

Онтология ОСД позволяет создать рекомендательную систему на её основе. Такая система имеет ряд преимуществ перед существующими подходами, одним из которых является простота внедрения в системы. Потенциально такой подход применим в программах, в которых необходимо реализовать систему подсказок, и системах продаж. Особыми техническими преимуществами являются: простота использования ОСД онтологии, учёт поведенческих факторов, простота расчётов, персонализированные рекомендации, малое количество данных для работы, учёт показа рекомендаций.

В результате аналитического обзора было обнаружено, что система, построенная на ОСД онтологии является гибридной. Она сочетает в себе разные возможности в зависимости от архитектуры сети. ОСД онтология может организовывать подход коллаборативная фильтрации item

based, делая прогноз на базе опыта, а также реализовать подходы основанные на контенте и основанные на знаниях. Такие возможности ОСД онтологии обусловлены из-за сетевой структуры и математической природы модели.

Всё выше перечисленное делает удобным применение технологии во многих технических и экономических системах.

#### Библиографический список

1. **Фальк, К.** Рекомендательные системы на практике [Текст] / К. Фальк. - Москва: ДМК Пресс, 2020. – 448 с.
2. **Бронфельд, Г.Б.** Инженерные основы моделей знаний : учеб. Н. Новгород: НГТУ им.Р.Е.Алексеева 2016. 139 с.
3. **Бронфельд, Г.Б.** Основы искусственного интеллекта : учеб. пособие. Н. Новгород: НГТУ им.Р.Е.Алексеева, 2014. – 253 с.
4. **Ландэ, Д. В.** Основы интеграции информационных потоков / Ландэ Д. В.//Информационные потоки.- 2006. - 240 с.
5. **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.
6. **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.
7. **Осовский, С.** Нейронные сети для обработки информации / Пер. с польского И.Д. Рудинского. – М.: Финансы и статистика, 2002. – 344 с.
8. **Тарков, М.С.** Нейрокомпьютерные системы: Учебное пособие / М.С. Тарков. – М.: Интернет-Университет Информационных Технологий; БИНОМ. Лаборатория знаний, 2006. – 142 с.
9. **Круглов, В.В., Борисов, В.В.** Искусственные нейронные сети. Теория и практика. – 2-е изд., стереотип. – М.: Горячая линия-Телеком, 2002–322 с.
10. **Ларичев, О.И.** Теория и методы принятия решений, а также Хроника событий в Волшебных Странах: Учебник. - М.: Логос, 2000. - 296 с : ил.
11. **Kenna, Ralph, Mac Carron, Pdraig.** Maths Meets Myths: Network Investigations of Ancient Narratives // Journal of Physics: Conference Series. — 2016. — Vol. 681. — P. 012002.
12. **Симонов, П. В.** Эмоциональный мозг. — М.: Наука, 1981. — 215 с.
13. Knowledge-based systems [Электронный ресурс]. – Режим доступа: [https://en.wikipedia.org/wiki/Knowledge-based\\_systems](https://en.wikipedia.org/wiki/Knowledge-based_systems) (дата обращения: 11.10.21)
14. Коллаборативная фильтрация для построения рекомендаций на основе данных о заказах [Электронный ресурс]. – Режим доступа: <https://cyberleninka.ru/article/n/kollaborativnaya-filtratsiya-dlya-postroeniya-rekomendatsiy-na-osnove-dannyh-o-zakazah/viewer> (дата обращения: 17.10.21)

**СИСТЕМЫ ИНТЕЛЛЕКТУАЛЬНОЙ ПОДДЕРЖКИ РАБОТЫ АВТОМАТИЗИРОВАННЫХ СИСТЕМ****П.С. Кулясов**

ORCID: 0000-0001-7793-2891 e-mail: p.kulyasov@gmail.com

**П.В. Мисевич**

ORCID: 0000-0002-7254-0079 e-mail: p\_misevich@mail.ru

**Е.Н. Панкратова**

ORCID: 0000-0002-5009-9315 e-mail: keibusan@gmail.com

**В.П. Хранилов**

ORCID: 0000-0003-1317-5320 e-mail: hranilov@nntu.ru

Нижегородский государственный технический университет им. Р.Е. Алексеева

В докладе описываются основные направления развития интеллектуализации автоматизированных систем. Предлагается классификационная система, которая состоит из двух уровней иерархии. На первом уровне позиционируются системы, которые основаны на аппарате искусственного интеллекта и ориентированы на решения задач в различных предметных областях. Развитие таких систем приводит к созданию программно-аппаратных комплексов – автоматизированных систем. Рост сложности инструментариев автоматизации решения задач в предметных областях приводит к необходимости создания специализированных систем, которые автоматизируют их создание и поддерживают в последующих этапах жизненного цикла. В рамках тренда интеллектуализации информационных технологий в основе этих инструментариев так же лежит аппарат искусственного интеллекта, и они решают задачи поддержки работы автоматизированных систем. Предлагается объединить эти инструментарии во второй – высший уровень иерархии классификационной системы. В качестве примера рассматривается динамическая модель функционирования систем второго уровня, которые построены на сценарно-ситуационном подходе.

**Ключевые слова:** автоматизированные системы, фреймовая модель, семантические сети, интеллектуальная поддержка, агентный подход.

**THE INTELLIGENT AUTOMATED SYSTEMS FOR SUPPORTING AUTOMATED SYSTEM****P.S. Kulyasov, P.V. Misevich, E.N. Pankratova, V.P. Khranilov**

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

The report describes the main trends of the IT intellectualization. A new two levels classification system is proposed. It is used for describing the intelligent system classes. The low level consolidates the intelligence systems for solving specific problems in various subject areas. The development of the systems leads to the creation of software and hardware complexes – automated systems. The increasing complexity of automation tools for solving problems in various subject areas leads to the development of specialized systems.

The intelligent specialized systems are situated on the high level of the classification system. At the level, the systems use artificial intelligence for supporting the automated systems during the life cycle. The design stage is supported by CAD systems. At the operational stage, automated systems support the processes in various subject areas. They are the following: monitoring systems, statistical collection systems, systems for supporting automated system operating.

The intellectual supporting systems for the resource allocation process are consolidated in a particular class in the high-level intelligent systems. The intelligent support systems for supporting scenarios in the automated system are consolidated in a separate class of intelligent support systems. The dynamic model of the functioning of the high-level systems is described. The model is used as an example of high-level intelligent systems.

**Keywords:** automated systems, frame model, semantic networks, intellectual support, agent system.

## Введение

Одним из глобальных трендов развития мира информационных технологий является интеллектуализация. Основными направлениями этого тренда являются:

- интеллектуальные энергосети;
- умный дом, район, город;
- интеллектуальные системы транспорта (беспилотные транспортные средства, интеллектуальные подсистемы технических систем и др.);
- диагностические системы в различных предметных областях (медицина, технические системы, экологические системы и др.);
- системы интеллектуальной поддержки (ИП) бизнес-процессов в различных предметных областях (биржевой трейдинг, управление процессами в организации, контроллинг, аудит и др.);
- системы поддержки принятия решений в различных предметных областях: экономика (центры ситуационного управления и другие диспетчерские центры);
- аналитические системы (анализ текстов, изображений, больших массивов данных, временных рядов и др.).

Приведенная выше классификация направлений интеллектуализации описывает низший слой иерархии инструментов, которые работают непосредственно с конкретной предметной областью.

## Анализ

Как правило, совокупность инструментальных решений задач в конкретной предметной области объединяются в автоматизированную систему, которая является объектом высокого уровня сложности и нуждается в поддержке. Поэтому на более высоком уровне иерархии расположим системы интеллектуальной поддержки функционирования автоматизированных систем, которые решают различные задачи (том числе и интеллектуальные) в предметных областях поддержки автоматизированных систем высокого уровня сложности.

Такие системы ИП описывают при помощи универсального аппарата, например [1]. Системы интеллектуальной поддержки данного класса строятся на различных принципах, но их объединяет общее: поддержка работы программно-аппаратного комплекса автоматизированных систем высокого уровня сложности и применение аппарата искусственного интеллекта.

**Определим систему интеллектуальной поддержки (СИП) автоматизированной системы (АС)** (по аналогии с системами интеллектуальной поддержки лиц принятия решений [2]) как комплекс, целью работы которого является интеллектуальная поддержка работы АС в течение жизненного цикла (ЖЦ). Поддержка заключается в том, чтобы помочь лицу-пользователю автоматизированной системы или подсистеме АС оценить ситуацию во внешней/внутренней среде (автоматизированной системы) и спланировать действия, востребованные этой ситуацией.

Атрибуты, определяющие современную систему интеллектуальной поддержки автоматизированной системы, следующие:

- в основе математического и алгоритмического обеспечения систем интеллектуальной поддержки лежит аппарат искусственного интеллекта;
- системы обеспечивают программно-аппаратным комплексам потребительские качества, которые дают разработкам конкурентные преимущества (при сравнении с аналогами).

Рассмотрим качества систем интеллектуальной поддержки автоматизированной системы несколько подробнее. В основе работы таких систем интеллектуальной поддержки лежит следующий аппарат:

- продукционные системы [3];
- системы фреймов [4];
- семантические сети [5];



- нечеткая логика [6-7];
- нейронные сети [8];
- гибридные формы представления знаний [9-12].

Рассмотрим информационное обеспечение систем интеллектуальной поддержки подробнее. Продукционные системы ориентированы на модификацию и расширение. Это определяется уникальным свойством правил продукций: высокий уровень независимости каждого элемента продукционной системы, и, как следствие, аддитивность и локальность каждого из изменений в системе. Продукционная система может начать работать по принципиально новой цепочке логического вывода, если изменить одно продукционное правило.

Эти свойства продукционных систем стали основой распространения инструментариев построения экспертных систем в форме скелетных оболочек. Построение подсистем интеллектуальной поддержки АС на базе продукционных систем ориентирует СИП на модификацию и расширение.

В системах интеллектуальной поддержки низкого уровня на базе продукционных систем, как правило, строят экспертные системы поддержки принятия решений в диспетчерских центрах (например, центры ситуационного управления (ЦСУ), центры управления перевозками и др.).

Использование продукционных систем в качестве информационного обеспечения интерпретаторов сценариев функционирования автоматизированных систем позволяет сделать сценарии гибкими и открывает возможность их приспособления к изменениям во внешней/внутренней среде АС.

Применение сетей фреймов [5] позволяет гармонично использовать идеи ситуационного подхода при построении систем поддержки принятия решений (ЦСУ и др.), систем мониторинга [13-14], систем интеллектуальной поддержки групп администратора в компаниях [14]. Фреймовый подход ориентирован на описание процессов в предметной области машинно-независимым способом – сетью фреймов, которая выступает, по сути, в роли концептуальной модели. Это позволяет с малыми затратами создавать инструментарии «трансформации» фреймовой концептуальной модели в информационное и программное обеспечение АС. Фреймовый подход хорошо себя зарекомендовал в разработках [15-17]

Семантическая сеть – статичная форма представления знаний, которая идеально подходит для составления системы стандартов в интероперабельных системах, в системах поддержки стандартов в промышленности и др.

Нейронные сети активно применяются для решения множества задач в различных предметных областях: систем мониторинга работы компьютерных сетей, интеллектуальной обработке данных, распознавания изображений, задачах анализа текстов и др. Тормозит активное развитие этого аппарата сложно формализуемая процедура выбора архитектуры нейронной сети – каузального фактора успеха всего проекта в целом.

Аппарат нечеткой логики активно применяется и для решения задач в конкретных предметных областях и для построения СИП АС в течение ЖЦ [16-17].

Гибридные базы знаний создаются, как правило, для нивелирования недостатков каждой формы представления знаний [10]. Например, они позволяют сочетать достоинства продукционных систем и сетей фреймов [18], достоинства фреймов и специфику применения мультимедийных фреймов [17], комбинировать нечеткую логику с ситуационным подходом [10].

Особый интерес представляет методология и технологии интерактивного управления ресурсами технических систем при проектировании [18], которая позволяет решать задачи интеллектуальной поддержки процессом управления ресурсами на базе моделей математического обеспечения СИП высшего уровня иерархии. Отметим, что комплекс задач управления ресурсами охватывает универсальную задачу подбора комплектующих в программно-аппаратный комплекс, которая является востребованной на этапе

проектирования систем различного уровня сложности [18] и распределения ресурсов в системах ИБ [19].

### Методика

Второй универсальной задачей АС является поддержка сценария функционирования АС в течение всего ЖЦ [20-21]. В ее основе лежит динамическая модель, которая при помощи сценарного подхода описывает процессы функционирования в АС [22]. Работу АС опишем системой:

$$\left\{ \begin{array}{ll} \left\{ \begin{array}{l} s_{k+1}^l = f(s_k^{\psi}, r_{j_k}^{\psi}) \\ s_{k+1}^l = \xi_k(\cdot), \end{array} \right. & \begin{array}{l} \text{если не сработал ни один агент} \\ \text{если сработал хотя бы один агент} \end{array} \end{array} \right. \quad (1)$$

$$\left\{ \begin{array}{l} r_{j_k}^{\psi} = \varphi(s_k^{\psi}, D_{system}^k) \end{array} \right. \quad (2)$$

$$\left\{ \begin{array}{l} D_{system}^{k+1} = C(s_{j_k}^{\psi}, D_{system}^k) \end{array} \right. \quad (3)$$

$$\left\{ \begin{array}{l} O_k = \lambda(s_k^{\psi}) \end{array} \right. \quad (4)$$

Где

- $f$  – функционал, который описывает переключение между событиями сценария;
- $s_k^{\psi}$  – событие сценария, которое выполняет АС в момент времени  $k$ ;
- $s_{k+1}^l$  – событие сценария, которое будет выполняться АС в момент времени  $k+1$ ;
- $r_{j_k}^{\psi}$  – идентификатор причинно-следственной связи события  $s_k^{\psi} \in S$ , по которой передается управление;
- $\varphi$  – теоретико-множественный функционал описания процесса формирования флага исхода  $r_{j_k}^{\psi}$  события  $s_k^{\psi}$ ;
- $C$  – функционал описания ситуации;
- $\lambda$  – функционал описания связь множества операций  $O_k$  с событием  $s_k^{\psi}$ ;
- $\xi_k(\cdot)$  – функция учета наличия механизма переключения с одного сценария на другой (ситуация при которой агент переключает работу АС с аварийного сценария на корректирующий).

Уравнение (1) описывает передачу управления с одного события сценария к другому по причинно-следственной связи. Флаг  $r_{j_k}^{\psi}$  исхода события (1) отображает результат анализа ситуации во внешней /внутренней среде АС. Уравнение (3) описывает последовательность описания ситуаций  $D_{system}^k$  моменты времени  $k$ .

Элементу  $s_k^{\psi} \in S$  соответствует множество операций  $O_k$ , которое описывается в модели соотношением (4).

Элемент  $\xi_k(\cdot)$  в системе представлен:

$$\left\{ \begin{array}{l} \xi_k(\cdot) = 0, \text{ переключения с одного сценария на другой не происходит (агент не} \\ \text{иницирован);} \\ \xi_k(\cdot) = s^{a_i}, \text{ где } s^{a_i} \text{ – событие корректирующего сценария, в которое система} \end{array} \right.$$

«переключилась».

Если  $\xi_k(\cdot) = 0$ , то (1) примет вид:

$$s_{k+1}^l = f(s_k^w, r_{jk}^w).$$

Если  $\xi_k(\cdot) = s^{a_i}$ , то (1) примет вид:

$$s_{k+1}^l = \xi_k(\cdot).$$

$a_i$  – это агентная процедура преодоления проблем сценария решения типовой задачи в предметной области (1). Совокупность элементов  $a_i$  образует множество  $A = (a_1, a_2, \dots)$  – множество агентов поддержки сценария в течение всего ЖЦ.

Функция  $a_i(s_k^w, D_k)$  принимает два значения:

$$\left\{ \begin{array}{l} a_i(s_k^w, D_k) = 0, \text{ если агент не сработал;} \\ a_i(s_k^w, D_k) = s^{a_i}, \text{ если агент сработал и информирует систему о необходимости} \\ \text{переключения сценария } \xi_k(\cdot) \neq 0. \end{array} \right.$$

Перепишем (1) в виде рекуррентного уравнения:

$$\left\{ \begin{array}{l} s_{k+1}^l = f(s_k^w, r_{jk}^w), \text{ если не сработал не один агент} \\ s_{k+1}^l = \xi_k(a_1, a_2, \dots), \text{ если сработал хотя бы один агент} \end{array} \right. \quad (5)$$

Если по результату ситуационного анализа ни один агент-элемент множества  $A$  не сработал, то мы осуществляем движение по сценарию верхнего уровня (1).

Если агент сработал, то управление передается сценарию более низкого уровня, который обозначается символом  $s^{a_i}$ . Коррекция завершается возвратом к событию сценария верхнего уровня с целью продолжения решения типовой задачи в предметной области функционирования АС.

Появление агентов является следствием описания сценариев работы АС иерархической моделью. На высшем уровне иерархии располагается сценарий решения АС задачи в предметной области. Если при работе АС возникли проблемы, то осуществляется «переключение» на корректирующий сценарий (сценарий более низкого уровня иерархии). Если этот корректирующий проблемы сценарий успешно завершен, то система возвращается на сценарий решения задачи в предметной области. Если при выполнении АС корректирующего сценария возникли проблемы, то для их преодоления формируем корректирующий сценарий более низкого уровня иерархии.

В [22] доказано, что модель (1)-(4) относится к классу динамических моделей, т.е. удовлетворяет системе аксиом Беллмана [23].

### Результаты

На базе динамической модели (1)-(4) разработаны системы интеллектуальной поддержки в течение их жизненного цикла, которые повышают устойчивость функционирования программно-аппаратных комплексов.

Таким образом, интеллектуализация, как тренд развития мира информационных технологий, имеет огромный потенциал в части поддержки автоматизированных систем в течение жизненного цикла. Видится перспективным использовать аппарат искусственного интеллекта для создания инструментариев поддержки программно-аппаратных комплексов, которые автоматизируют все этапы жизненного цикла и реализованы как в форме автономных комплексов (САПР, САД и др), подключаемых (при необходимости) комплексов так и в форме инструментариев сопровождения систем, которые реализованы в форме отдельных подсистем или комплексов.

### Библиографический список

1. **Khranilov, V.P.** Models in a State Space for Engineering Applications / V.P. Khranilov, V.V. Andreev, P.V. Misevich, E.N. Pankratova // 2019 XXI International Conference Complex Systems: Control and Modeling Problems (CSCMP). Pp. 788-792.
2. **Басалин, П.Д.** Модели и методы интеллектуальной поддержки процессов принятия решений. Учебное пособие / П.Д. Басалин, К.В. Безрук, М.В. Радаева // Нижний Новгород: Нижегородский госуниверситет. 2011. 108 с.
3. **Гаврилова, Т.А.** Базы знаний интеллектуальных систем. Учебник / Т. А. Гаврилова, В. Ф. Хорошевский // СПб.: Питер. 2000.
4. **Roussopoulos, N.D.** A semantic network model of data bases / N.D. Roussopoulos // TR No 104, Department of Computer Science, University of Toronto. 1976.
5. **Minsky, M.** A framework for representing knowledge. / M. Minsky // MIT AI Laboratory Memo 306, June. 1974.
6. **Zadeh, L. A.** Fuzzy sets / L. A. Zadeh // Information and Control. Vol. 8. 1965. Pp. 338-353.
7. **Zadeh, L. A.** Fuzzy Sets as a Basis for a Theory of Possibility / L. A. Zadeh // Fuzzy Sets and Systems. Vol. 1. No. 1. 1978. Pp. 3-28.
8. **Хайкин, С.** Нейронные сети. Полный курс 2-е изд. Пер. с англ. / С. Хайкин // М.: Вильямс. 2006. 1104 с.
9. **Лазарев, В.А.** Модели информационного обеспечения систем поддержки работы комплексов автоматизированного тестирования программного обеспечения / В.А. Лазарев, П.В. Мисевич // Системы управления и информационные технологии. 2016. №4 (65). С. 44-48.
10. **Khranilov, V.P.** The Use of Hybrid Knowledge Bases in Designing Engineering Systems / V.P. Khranilov; P.V. Misevich; E.N. Pankratova // 2020 2nd International Conference on Control Systems, Mathematical Modeling, Automation and Energy Efficiency (SUMMA).
11. **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.
12. **Ермилов, А. Э.** Построение инструментальных систем мониторинга с использованием фрейм-модели и нечеткой логики / А. Э. Ермилов, П. В. Мисевич // Системы управления и информационные технологии. 2014. №3(57). С. 57-60.
13. **Ermilov, A.E.** Monitoring systems development using situational approach and fuzzy logic / A.E. Ermilov // 2019 IOP Conf. Ser.: Mater. Sci. Eng. 695 012016.
14. **Лазарев, В.А.** Применение производственной системы для сопровождения автоматического тестирования / В.А. Лазарев // Системы управления и информационные технологии. 2013. №4 (54). С. 98-103.
15. **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.
16. **Мисевич, П.В.** Использование логистического подхода к вопросам построения процедур выделения и преодоления аварийных ситуаций в автоматизированных системах / П.В. Мисевич // Системы управления и информационные технологии. 2006. N4.2(26). С. 256-261.
17. **Басалин, П.Д.** Организация процесса проектирования в САПР с использованием фреймово-производственной модели / П.Д. Басалин, П.В. Мисевич // Математическое моделирование и оптимальное управление: межвуз. тематич. сб. научн. тр. под ред. Р.Г. Стронгина. Н. Новгород: изд-во Нижегородского ун-та, 1996. С. 164-171.
18. **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. 189-193.
19. **Хранилов, В.П.** Интерактивное распределение доступа пользователей к программным модулям вычислительной сети / В.П. Хранилов, Д.В. Прохоров // Вопросы защиты

информации. 2007. №2 (77). С.38-44.

20. **Мисевич, П.В.** Логистический подход к вопросам проектирования систем объектов и формирования сценариев работы программного обеспечения / П.В. Мисевич // Системы управления и информационные технологии. 2007. № 4.1(30). С. 169-174.
21. **Мисевич, П.В.** Сценарно-ситуационный подход к проектированию средств интеллектуальной поддержки процесса функционирования автоматизированных систем / П.В. Мисевич // Системы управления и информационные технологии. 2007. N2.1(28). - С. 166-171.
22. **Мисевич, П.В.** Динамическая модель функционирования автоматизированной системы / П.В. Мисевич // Системы управления и информационные технологии. 2008. 3.1(33). С. 175-179.
23. **Kalman, R.E.** Topics In Mathematical System Theory / R. E. Kalman, P. L. Falb, M. A. Arbib // New York, Sydney. 1969.

## АНАЛИЗ МЕТОДОВ И АЛГОРИТМОВ ДИАРИЗАЦИИ В АУДИОАНАЛИТИКЕ

Т.В. Львутина<sup>1</sup>

ORCID:0000-0002-2061-8858 e-mail:tat.lvutina@mail.ru

К.И. Абросимов<sup>2</sup>

ORCID:0000-0001-9262-0474e-mail:abrosimov.kirill.1999@mail.ru

А.С. Суркова<sup>1</sup>

ORCID:0000-0003-0018-9053 e-mail:ansurkova@yandex.ru

<sup>1</sup>Нижегородский государственный технический университет им. Р.Е. Алексеева<sup>2</sup>Национальный исследовательский университет ИТМО

В докладе проводится количественный и качественный анализ, а также экспериментальное сравнение существующих алгоритмов диаризации с целью определения их достоинств и недостатков. Диаризация представляет из себя процесс разделения входного аудиосигнала на сегменты в соответствии с определением личности диктора. Для проведения экспериментов реализованы два метода детектора речевой активности (VoiceActivityDetection, VAD): метод, основанный на кратковременных (short-term) признаках и метод на основе MFCC (Mel-frequencySpectralCoefficients). В рамках векторного представления использована модель гауссовых смесей, обученная на относительно большом количестве голосовых данных (GaussianMixtureModelUniversalBackgroundModel, GMM-UBM) и алгоритм MAP адаптации (Maximum A-PosterioriAdaptation), векторное представление на основе частоты основного тона, эмбединги диктора на основе LSTM. Для кластеризации применен алгоритм спектральной кластеризации. Для оценки качества решения задачи диаризации использована метрика DER (DiarizationErrorRate). Данная метрика показывает, какая часть аудиофрагмента определена неверно. Эксперименты проводились с аудиофрагментами, в которых отсутствует и присутствует шумление, а также с аудиофрагментами, которые содержат реальные разговоры людей.

**Ключевые слова:** диаризация, алгоритмы VAD, mapadaptation, GMM, частота основного тона, embedding диктора, LSTM.

## ANALYSIS OF DIARIZATION METHODS AND ALGORITHMS IN AUDIO ANALYTICS

T.V. Lvutina<sup>1</sup>, K.I. Abrosimov<sup>2</sup>, A.S. Surkova<sup>1</sup><sup>1</sup> Nizhny Novgorod State Technical University n.a. R.E. Alekseev<sup>2</sup>National Research University ITMO

**Purpose:** In this article, an experimental comparison of diarization algorithms is made.

**Design / methodology / approach:** Two VAD methods are implemented for conducting experiments – based on short-term signs and based on MFCC. Within the framework of vector representation - GMM-UBM and the map adaptation algorithm, vector representation based on the pitch frequency, speaker embeddings based on LSTM. The spectral clustering algorithm is used for clustering. To assess the quality of solving the diarization problem, the DER metric (Diarization Error Rate) is used. Testing is performed with an audio fragment in which there is no noise, as well as with audio fragments that contain a real conversation of people.

**Findings:** Performing an experiment on an audio file that contains a real conversation of people, it turned out that GM and map adaptation show worse results than embedding based on STM.

**Research limitations/implications:** The present article provides a starting-point for further research in the field of diarization.

**Originality/value:** The value of the article lies in the experimental comparison of diarization algorithms.

**Key words:** diarization, VAD algorithms, mapadaptation, GM, pitch frequency, speaker embedding, LSTM.

## Введение

С каждым годом задача диаризации становится все более актуальной. Наиболее часто она используется в системах распознавания речи, в которых есть необходимость четкого разделения входного аудиосигнала на сегменты, принадлежащие разным дикторам. Существует большое количество методов, позволяющих произвести диаризацию. Различная сложность этих методов обуславливается многообразием подходов к определению личности дикторов. Некоторые методы показывают хорошие результаты при выполнении задачи диаризации даже при наличии шума в аудиофайле, однако такие алгоритмы имеют высокую сложность, требуют большие вычислительные мощности и временные затраты. Основная задача данной работы состоит в анализе существующих подходов и выборе оптимальных для практического использования алгоритмов диаризации, в зависимости от дополнительных условий проблемы.

## Постановка задачи диаризации

Диаризация представляет из себя процесс разделения входного аудиосигнала на сегменты в соответствии с определением личности диктора. Цель диаризации – ответить на вопрос «когда и кто говорит?». Этот процесс сочетает в себе задачи сегментации и кластеризации говорящего. Сегментация направлена на поиск точек смены диктора в аудиосигнале. Кластеризация служит для группирования речевых сегментов на основе характеристик говорящего.

Процесс диаризации можно представить в виде трех последовательных блоков:

1. Обнаружение речи в системах распознавания диктора выполняет **детектор речевой активности** (Voice Activity Detection, VAD) — алгоритм, классифицирующий исходные участки аудиосигнала как речь или не речь.
2. **Сегментация** — это задача нахождения в аудиосигнале точек, в которых произошла смена диктора. Методы сегментации можно разделить на две основные категории:
  - *Метрическая*: определяет, исходят ли два акустических сегмента из одного динамика, вычисляя расстояние между двумя акустическими сегментами. Сегментация на основе метрик является наиболее популярным типом метода, так как никаких предварительных знаний не требуется для ее работы.
  - *Модельная*: модели с помощью учителя обучаются распознавать дикторов, а затем используются, для оценивания, где есть точки изменения в аудиофайле.
3. При **кластеризации** сегментов изучается сходство между ними, и строится иерархия кластеров. Существуют два основных подхода к созданию такой иерархии:
  - *Агломеративный*: это подход снизу-вверх (англ. bottom-up). Первоначально для каждого сегмента существует один кластер, и кластеры итеративно объединяются, пока набор кластеров не достигнет оптимального размера.
  - *Разделяющий*: это подход сверху-вниз (англ. top-down). В начале все сегменты хранятся в одном большом кластере. Новый диктор вводится рекурсивно, и кластер разбивается на все меньшие и меньшие кластеры.

## VAD: методы и алгоритмы

Аудиосигнал представляет из себя поочередно сменяющиеся интервалы информационно нагруженных сегментов и пауз. Интервалы, в которых присутствует голосовая информация, называют участками активной речи, а паузы между речевыми «всплесками» – участками пауз. Когда система детектирует участки пауз, передача, как правило, прекращается и передается только общее описание информации о фоне. VAD основное внимание уделяет следующим особенностям:

- речь является нестационарным процессом;
- фоновый шум стационарен на более длинном отрезке времени по сравнению с речью;
- уровень речевого сигнала обычно выше уровня фонового шума.

Наиболее сложным оказывается справиться с устойчивостью к шуму. В условиях высокого SNR (Signal-to-noiseratio), простейшие VAD алгоритмы работают удовлетворительно, но при условиях низкого SNR многие алгоритмы VAD показывают неудовлетворительные результаты. Также следует отметить, что детектор речевой активности должен быть достаточно простым, для возможности его использования в реальном времени. Таким образом, основными характеристиками VAD являются устойчивость и простота. Эти характеристики учитываются при выборе алгоритма [3].

Из большого числа предложенных алгоритмов VAD чаще всего используются методы Short-term Energy и zero-crossingrate из-за своей простоты. Однако, они показывают неудовлетворительные результаты при наличии шумов. Для того что бы исправить этот недостаток были предложены разные устойчивые акустические характеристики на основе — функции автокорреляции, спектра(spectrumbased), мощности на узкополосном отрезке (power in the band-limited region), MFCC (Mel-frequency Cepstral Coefficients — кепстральные коэффициенты тональной частоты), дельт спектральных частот (delta line spectral frequencies) и статистик высшего порядка. Эксперименты показали, что использование этих характеристик приводит к увеличению устойчивости VAD к шумам [1].

В работе использовались следующие два алгоритма VAD. Первый на основе кратковременных (short-term) признаков [5], который весьма устойчив к шуму, и опирается на такие параметры сигнала, как краткосрочная энергия, мера спектральной плоскостности и преобладающие частоты в анализируемом интервале аудиозаписи. Второй алгоритм VAD – пороговый, на основе MFCC [3], то есть из аудиозаписи извлекаются MFCC и на основе первых трех коэффициентов и заданного порогового значения принимается решение о наличии голосовой активности на анализируемом интервале.

### Гауссовы смеси в задаче diarизации и алгоритм MAPAdaptation

После проведения процедуры VAD имеем отдельные сегменты входного аудиосигнала, представляющие из себя речь и паузы. На следующем этапе необходимо смоделировать каждого диктора. Для этого следует извлечь признаки из каждого сегмента аудиосигнала и на этих данных обучить модель. Этот этап можно выполнить с помощью модели гауссовых смесей (Gauss Mixture Model, GMM).

Данная модель представляет собой взвешенную сумму Гауссиан:

$$p(x|\lambda) = \sum_{i=1}^M w_i p_i(x),$$

где  $\lambda$  – модель диктора,  $M$  – количество компонентов модели,  $w_i$  – веса компонентов. Веса компонентов подбираются таким образом, чтобы их сумма была равна единице:  $\sum_{i=1}^M w_i = 1$ .

Функция плотности вероятности каждого компонента задаётся формулой:

$$p_i(x) = \frac{1}{2\pi^{\frac{D}{2}} \sqrt{\Sigma}^{\frac{1}{2}}} \exp\left(-\frac{1}{2}(x - \mu_i)^T \Sigma^{-1} (x - \mu_i)\right),$$

где  $D$  – размерность пространства признаков,  $\mu_i$  – вектор математического ожидания;  $\Sigma$  – матрица ковариации.

Можно ожидать, что модели, обученные на сегментах, принадлежащих одному и тому же человеку, будут схожи. Однако, на этом этапе мы сталкиваемся со следующей проблемой: при извлечении признаков из аудиофайла длиной 1 сек с частотой дискретизации 8000 Гц при размере окна 10 мс, получаем набор из 800 векторов MFCC. Этим данным не хватит для обучения модели. Решением этой проблемы может стать использование Универсальной Фоновой Модели (Universal Background Model, UBM), её еще называют дикторонезависимой. Идея использования UBM заключается в следующем. При обучении



GMM на большой выборке данных получаем на выходе акустическую модель обобщённого диктора (это и будем называть UBM). А затем, с помощью алгоритма адаптации Maximum A-Posteriori Adaptation (MAP Adaptation), обобщенная модель адаптируется под признаки, извлекаемые из каждого сегмента.

В общем случае алгоритм Maximum A-Posteriori Adaptation можно представить в виде следующих шагов.

**Шаг1.** Рассчитать апостериорную вероятность на адапционных данных и достаточные статистики для веса, медианы и дисперсии каждой гауссианы.

**Шаг2.** Затем полученные статистики комбинируются с параметрами UBM и получаются параметры адаптированной модели.

В основе алгоритм MAP адаптации лежат следующие вычисления:

$$E_i = \frac{1}{n_i} \sum_{t=1}^N \Pr(\hat{\mu}_i | x_t) x_t$$

$$\hat{\mu}_i = \beta_i E_i + (1 - \beta_i) \mu_i$$

$$\beta_i = \frac{n_i}{n_i + r^\beta}$$

Здесь  $\Pr(\hat{\mu}_i | x_t)$  — апостериорная вероятность,  $E_i$  — достаточная статистика для  $\mu$ ,  $\hat{\mu}_i$  — медиана адаптированной модели,  $\beta_i$  — коэффициент адаптации,  $r^\beta$  — фактор соответствия.

В начале происходит обучение UBM (двухкомпонентная смесь гауссовых распределений) на входных данных. В результате этого появляются новые данные, которые не укладываются в модель. Алгоритм MAP адаптации служит для смещения центровгауссиан так, чтобы они ложились на новые данные. Применяя этот алгоритм на экспериментальных данных, ожидается, что на сегментах с одним и тем же диктором гауссианы будут смещаться в одном направлении, образуя таким образом кластеры. Именно поэтому для кластеризации сегментов будут использоваться данные о сдвиге  $\mu$ .

### Диаризация на основе частоты основного тона

*Частота основного тона*, Fundamental Frequency или F0 – это частота колебания голосовых связок при произнесении тоновых звуков. При произнесении нетоновых звуков, например говорении шепотом или произнесении шипящих и свистящих звуков, связи не колеблются, а значит эта характеристика для них не релевантна.

Вариабельность частоты основного тона довольно велика, причем она может сильно отличаться не только между людьми (для более низких в среднем мужских голосов частота составляет 70-200 Гц, а для женских может достигать 400 Гц), но и для одного человека, особенно в эмоциональной речи.

Определение частоты основного тона применяется для решения широкого спектра задач:

- Распознавание эмоций;
- Определение пола;
- При решении задачи сегментации аудио с несколькими голосами или разделения речи на фразы.

Методы определения F0 можно разделить на три категории:

- основанные на временной динамике сигнала или time-domain;
- основанные на частотной структуре или frequency-domain;
- комбинированные методы.

Комбинированные методы определения F0, сочетают в себе вычисление как частотной, так и временной информации. В работе использован алгоритм YAAPT — Yet Another Algorithm of Pitch Tracking. Рассмотрим этапы работы данного алгоритма.

*1 этап.* Первым этапом алгоритма YAAPT является преобращение. На этом этапе значения изначального сигнала возводят в квадрат, таким образом формируется вторая версия сигнала. Этот шаг служит для усиления и восстановления “затертых” пиков автокорреляции.

*2 этап.* По спектру преобразованного сигнала рассчитывается базовая траектория F0. Кандидаты на F0 определяются с помощью функции SpectralHarmonicsCorrelation (SHC). Также по спектральной мощности происходит определение фреймов voiced-unvoiced, после чего ищется наиболее оптимальная траектория, при этом учитывается возможность pitch doubling/pitch halving.

*3 этап.* Как для изначального сигнала, так и для преобразованного производится определение кандидатов на F0, и вместо автокорреляционной функции здесь используется Normalized Cross Correlation (NCCF).

*4 этап.* Производится оценка всех возможных кандидатов и вычисление их значимости, или веса (merit). Вес кандидатов, полученных по аудио сигналу, зависит не только от амплитуды пика NCCF, но и от их близости к траектории F0, определенной по спектру. То есть частотный домен считается хоть и грубым в плане точности, но зато устойчивым.

*5 этап.* Для всех пар оставшихся кандидатов рассчитывается матрица Transition Cost — цены перехода, по которой в итоге и находят оптимальную траекторию.

В итоге, после полученной траектории частоты основного тона на анализируемом интервале, значения сортируются в порядке возрастания и для векторного представления берутся 15 самых малых значений и 15 самых больших значений, то есть 30-мерный вектор.

### **Диаризация на основе извлечения embedding диктора**

Сжатые векторные представления (Embeddings, эмбеддинги) – современное направление векторного представления объектов сложной структуры. Впервые были применены к векторному представлению текстовой информации (модели Word2Vec). Для получения эмбеддингов, как правило, используют различные архитектуры нейронных сетей. Яркими представителями моделей эмбеддингов являются автоэнкодеры.

В области речевых технологий эмбеддинги диктора используются в таких задачах, как диаризация, синтез речи, распознавание диктора. Для получения эмбеддингов добавляется дополнительный кодировщик при решении задач, представленных выше. Например, модель Resemblyzer [6] использует нейронную сеть кодировщик на основе LSTM. Данная модель из полученного аудио фрагмента извлекает несколько 256-мерных векторов по нескольким подфрагментам, после чего производит их усреднение. Полученный вектор и является эмбеддингом диктора. Стоит отметить, что модель не определяет – говорит ли в этом фрагменте один диктор или несколько, модель лишь получает вектора, усредняет и производит L2-нормализацию. В экспериментах используется именно Resemblyzer предобученная модель.

### **Экспериментальное сравнение методов диаризации**

В рамках эксперимента рассматривается два метода VAD– на основе short-term признаках и на основе MFCC. В рамках векторного представления – GMM-UBM и алгоритм map adaptation, векторное представление на основе частоты основного тона, эмбеддинги диктора на основе LSTM. Для кластеризации используется алгоритм спектральной кластеризации.

Для оценки качества решения задачи диаризации используется метрика DER (Diarization Error Rate) [7]. Данная метрика показывает, какая часть аудиофрагмента определена неверно.

### Оценка качества при явных паузах между фразами дикторов

Первый эксперимент – применение методов диаризации к аудиофрагментам, в которых два диктора между фразами друг друга выдерживают явную паузу. Для тестирования используется 5 минут размеченных аудиофрагментов. Стоит отметить, что в фрагментах почти нет шума.

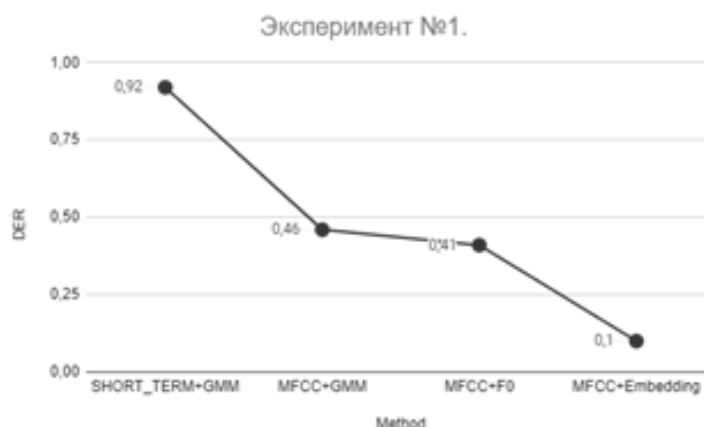


Рис. 1. Эксперимент №1

На рис.1 представлен график полученных результатов. Стоит отметить, что VAD на основе short-term признаков показал себя крайне плохо, поэтому в дальнейшем использовался только VAD на основе MFCC. Гауссовы смеси с алгоритмом map adaptation показали примерно одинаковый результат, 46% и 41% ошибок соответственно. Наилучший результат получился у метода диаризации на основе эмбеддингов диктора, полученные из предобученной нейронной сети LSTM. На рис.2 представлен пример ожидаемого разделения аудиофрагмента по дикторам и полученного с помощью различных методов.

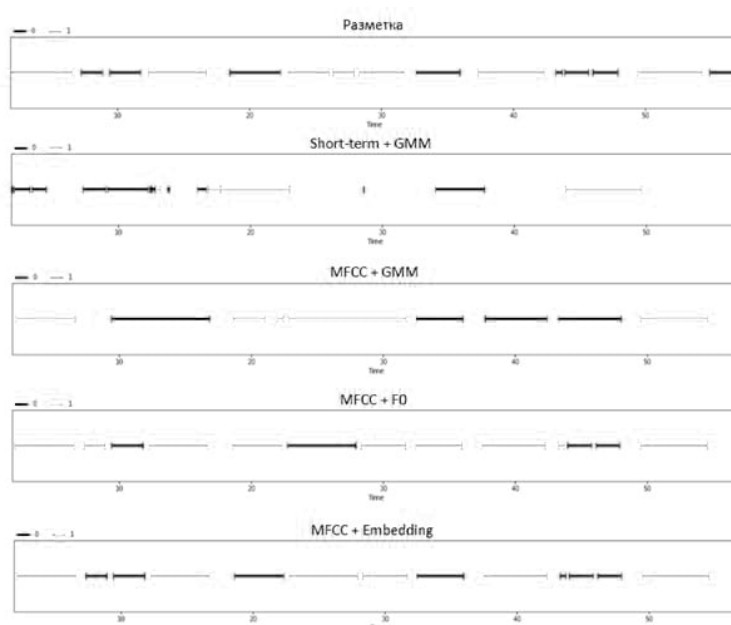


Рис. 2. Разделение аудиофрагмента по дикторам

### Оценка качества при спорах и перебиваниях

Далее был проведен более сложный эксперимент, для которого было взято 5 минут реального разговора людей и размечено на дикторов. В разговоре люди перебивали друг

друга, поддакивали, к тому же качество записи немного хуже, чем в первом эксперименте. Результаты эксперимента приведены на рис.3.



Рис. 3. Эксперимент №2

Как можем заметить, результаты хуже, чем в первом эксперименте. Однако, тенденция качества также видна, что GMM и adaptation явно хуже себя показывает, чем эмбединги на основе LSTM. На рис.4 представлен пример размеченного фрагмента и предсказание MFCC+Embeddings.

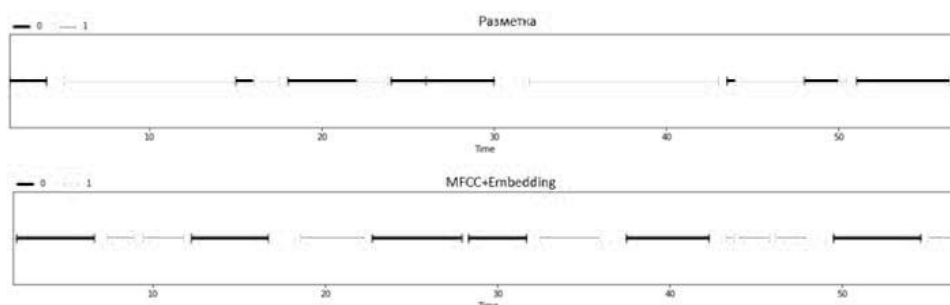


Рис. 4. Пример размеченного фрагмента

### Оценка качества при зашумленном аудиофрагменте.

В этом эксперименте для тестирования используются те же данные, что и в первом, однако теперь во все аудиоданные добавлен белый шум. На рис.5 представлены результаты эксперимента.

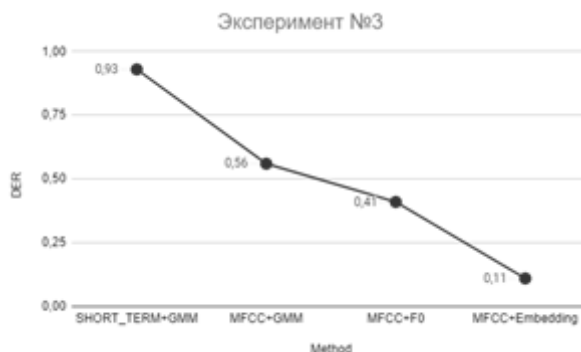


Рис. 5. Эксперимент №3

Стоит отметить, что результаты в целом не изменились, однако шум серьезно повлиял на результаты метода GMM с map adaptation.

## Заключение

В работе была произведена реализация различных методов диаризации. Реализованы два алгоритма детектирования голосовой активности – VAD на краткосрочных признаках показал себя очень слабо, однако VAD на основе MFCC с порогового значения по первым трем фильтрам на чистых данных показал себя достойно и существенно внес вклад в качество решения задачи диаризации. Было выяснено, что наилучшим векторным представлением диктора в аудио фрагментах являются эмбединги, получаемые из рекуррентной нейронной сети. Весьма неплохо себя показало векторное представление, на основе частоты основного тона, однако, это решение не всегда применимо, потому что частоты основного тона у нескольких людей в аудиозаписи могут быть близки, а следовательно, и при кластеризации полученных векторов, они могут быть объединены в один кластер, а значит и иметь метку одного и того же диктора.

Диаризация – это сложная задача аудиоаналитики, которая активно используется в распознавании речи на дальнем микрофоне. Дальнейшими перспективными исследованиями являются разработка и анализ качества эмбедингов диктора на основе различных архитектур глубоких нейронных сетей. Также в связи с прорывами в области NLP, применяют распознавание речи с дальнейшим анализом контекста для поиска точек смены диктора.

## Библиографический список

1. **Prasad, R.** Comparison of Voice Activity Detection Algorithms for VoIP [Текст]/R. Prasad [et al.]/In proc. 7th IEEE symp. on Computers and Communications (ISCC'02).–2002.–P. 530.
2. **Борисевич Д.А., Давыдов, Г.В., Попов, В.А.** Обнаружение речи в сигналах в режиме реального времени // Доклады БГУИР. – 2013. – № 5 (75). – С. 100–103.
3. **Зельманский, О.Б., Давыдов, А.Г.** Построение программного модуля сегментации речи на основе анализа изменения спектра // Безопасность информационных технологий. – 2011. – № 3. – С. 127–133.
4. **Beritelli, F.** Performance Evaluation and Comparison of G.729/AMR/Fuzzy Voice Activity Detectors. [Текст]/F. Beritelli [et al.]/IEEE Signal Proc. Lett.–2002.–Vol. 9.–№ 3
5. **Moattar, M. H., Homayounpour, M. M.** A Simple but efficient real-time Voice Activity Detection Algorithm / 17th European Signal Processing Conference (EUSIPCO 2009) / Glasgow, Scotland, August 24-28, 2009 – P. 2549–2553.
6. **Li, Wan, Quan, Wang, Alan, Papir, Ignacio, Lopez Moren.** Generalized End-to-End loss for speaker Verification. — 2020. — Vol. arXiv: 1710.10467. — version 5. Access mode: <https://arxiv.org/pdf/1710.10467.pdf>
7. **Bagga, A. and Baldwin, B.** Algorithms for scoring coreference chains. – 1998. – Proceedings of LREC 1998.

## ГРАФИЧЕСКОЕ ПРЕДСТАВЛЕНИЕ СИНТАКСИЧЕСКИХ КОНСТРУКЦИЙ В АСПЕКТЕ ФУНКЦИОНАЛЬНО-СТИЛИСТИЧЕСКОЙ ДИАГНОСТИКИ (ПРОСТОЕ ПРЕДЛОЖЕНИЕ)

А.В. Синелева

ORCID: 0000-0001-9926-9760 e-mail: [sinstasi@mail.ru](mailto:sinstasi@mail.ru)

Нижегородский государственный университет им. Н.И. Лобачевского

В докладе представлен фрагмент алгоритма моделирования простых предложений и их графического представления в аспекте определения параметров простого предложения, которые должны рассматриваться при описании авторского и функционального стиля и которые могут выступать в качестве дифференцирующих. Метод интерпретации синтаксических единиц, репрезентирующих предложение, разработанный проф. И.П. Севбо, показывает возможности данного метода анализа и описания синтаксических конструкций. Рассматривается необходимость модернизации некоторых соглашений теории графического представления синтаксических структур, предложенной И.П. Севбо, связанная с изменением основного положения: отражения в схеме не слов, а синтаксических позиций. Это позволяет, с одной стороны, несколько упростить схему, а с другой, дает возможность отражать важнейшие типологические характеристики простого предложения.

**Ключевые слова:** граф, деревья зависимостей, монопредикат, односоставные предложения, простое предложение, синтаксическая категория, синтаксическая позиция, функциональный стиль.

## GRAPHIC REPRESENTATION OF SYNTAX CONSTRUCTIONS IN THE ASPECT OF FUNCTIONAL-STYLISTIC DIAGNOSIS (SIMPLE SENTENCE)

A.V. Sineleva

Nizhniy Novgorod state university n. a. N.I. Lobachevsky

**Purpose:** This article considers the need to make adjustments to the algorithm for graphical representation of syntactic structures, and proposes a change in the basic provision regarding units of the taxonomic level.

**Design/methodology/approach:** Refinements of the algorithm of models of syntactic structures in the form of a dependency grammar developed by I.P. Sevbo are proposed. One of the basic principles for constructing a grammar of dependencies is the provision that a sentence is built from units of a taxonomic level - word forms, while functional words are considered on a par with significant parts of speech. The article presents clarifications, primarily related to the change in units of the taxonomic level, as which it is proposed to consider not individual words, but syntactic positions, since the units of the syntactic level of the language are syntactic categories. In addition, it is proposed to remove prepositions from schemes that complicate constructions: a preposition constitutes a word form, and not a separate syntactic position, and graph nodes must correspond to syntactic positions.

**Findings:** A refined approach to the interpretation of syntactic structures is proposed. The new provisions make it possible to differentiate and describe the structural features of a simple sentence: a one-part sentence or a two-part sentence in terms of the severity of grammatical basis, type of a monopredicate, the presence of incompleteness, which are reflected in corresponding sentence models.

**Research limitations/implications:** This article contains clarifications that relate to the structure of simple uncomplicated sentences and simple sentences complicated by homogeneous members. Proposals that have such structural complications as separate definitions, circumstances and additions, introductory, plug-in constructions will be considered later.

**Originality/value:** The modeling of simple sentences, obtained as a result of applying a modified interpretation algorithm, is a continuation of the development of a structural direction in linguistics. Graphs built according to the refined algorithm can be used both for stylistic diagnostics and for the formation of a linguistic hypothesis regarding the features of the individual author's style.

**Keywords:** graph, dependency trees, monopredicate, one-part sentences, simple sentence, syntactic category, syntactic position, functional style.

## Введение

В отечественной лингвистике достаточно давно предпринимались попытки графического представления предложений: А.М. Пешковским были определены некоторые положения, приведшие впоследствии к созданию теории описания синтаксических единиц. В 70-80-е годы XX века профессором И.П. Севбо для анализа синтаксических особенностей авторских стилей был разработан метод представления и описания синтаксических конструкций (простых и сложных предложений) в виде деревьев зависимостей [4]. Основные принципы построения грамматики зависимостей сведены к следующим положениям:

- предложение строится из единиц таксономического уровня – словоформ, при этом служебные слова рассматриваются наравне со знаменательными частями речи;
- слова в предложении связаны отношением подчинения (согласования, управления, примыкания);
- в предложении есть одно абсолютно независимое слово – сказуемое, следствием чего является наличие двух типов связей: ветвлений и цепочек подчинений;
- для каждого слова задается набор сильных управлений, отражающих его поведение в речи при построении конкретных предложений;
- иерархическая структура зависимостей между словами предложения изображается в виде дерева зависимостей [4, с. 8-9].

Продолжая разработку моделирования синтаксических структур, предложенную профессором И.П. Севбо, мы говорили о необходимости некоторых уточнений, введение которых обусловлено снижением сложности изображения, во-первых, и возможности отражать в дереве зависимостей характеристики предложений (пока речь идет о простых предложениях), связанные их типологией, в частности, это структура грамматической основы и предиката, полнота или неполнота конструкции. Ранее нами было предложено уточнить некоторые положения, в первую очередь связанные с изменением единиц таксономического уровня, в качестве которых предлагается рассматривать не отдельные слова, а синтаксические позиции, т.к. единицами синтаксиса являются синтаксические категории, и, по мнению Б.Н. Головина, синтаксис определяется двумя категориями: синтаксическими категориями словесных позиций и синтаксическими категориями словесных конструкций. Кроме этого, из схем было предложено убрать предлоги, усложняющие конструкции, к тому же предлог составляет словоформу, а не отдельную синтаксическую позицию, ставить в соответствие фразеологизмам, которые занимают промежуточное положение между словом и словосочетанием и достаточно часто допускают замену одним словом, один узел – это одна синтаксическая позиция [6, с.858].

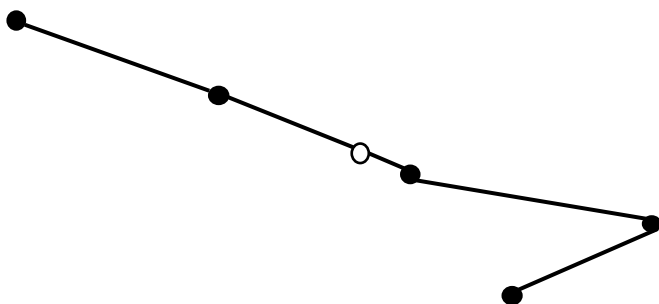
### Моделирование синтаксических структур в виде деревьев зависимостей

Как известно, основной единицей синтаксиса является предложение, потому что «в предложении находят выражение наиболее существенные функции языка: мыслительная (познавательная или экспрессивная) и коммуникативная... Предложение - самая многоаспектная единица синтаксиса, оно характеризуется совокупностью признаков, которые можно разделить на две группы: структурные предложение имеет специфические структурные схемы и особые способы выражения строевых элементов схемы и семантические (имеет модально-временной план, тема-рематическое членение и др.)» [1, с.161-162], «Предложение образовано по специальному абстрактному образцу, обладает своими языковыми значениями, формальными характеристиками. Простое предложение имеет разветвленный круг функций [3, с.7]. Предложение является очень сложной многогранной единицей, описывать которую можно в разных аспектах и направлениях: со стороны коммуникативной организации, логико-психологической, стилистической, структурно-семантической и др.

В данной статье продолжается рассмотрение единиц таксономического уровня, касающихся в основном структурно-грамматических характеристик простого предложения, а именно, односоставных предложений, являющихся важнейшими функционально-стилистическими характеристиками, а иногда и индивидуально-авторскими в художественном или публицистическом стиле (информационные анонсы, авторские аналитические обзоры, комментарии).

Конечно, частотность односоставных предложений в речи ниже, чем двусоставных, однако в текстах разной функционально-стилистической направленности представлены все типы односоставных предложений. Например, выборка объемом 200 предложений из романа В. Набокова «Подвиг» включает в себя 429 двусоставных предложений (самостоятельных и в составе сложных) и 73 односоставных, что составляет 14,5% от общего числа [5, с.56]. В выборке из 900 информационных анонсов стартовой страницы поисковой системы Яндекс за 2016-2018 односоставные предложения составляют 23,4% от общего числа проанализированных предложений [7, с.85]. Все виды односоставных предложений, как правило, включаются в список параметров при описании индивидуально-авторского стиля, так как они могут оказаться дифференцирующими.

Приведем в качестве примера модель простого предложения, построенного в соответствии с положениями и соглашениями, принятыми в теории И.П. Севбо, и рассмотрим возможность идентификации простых предложений по дереву зависимостей: корень дерева – сказуемое, незакрашенный узел соответствует предлогу – система соглашений и обозначений И.П.Севбо (рис.1).



*Раскапывайте погребенных в земле слепых исполинов (P1) [4, с.31]*  
*Вернулся я в город моего детства (P2)*  
*Вернулся давно в город моего детства (P3)*

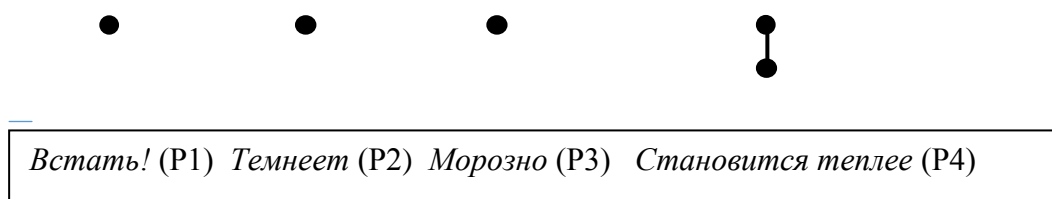
**Рис. 1. Структурная схема простого предложения**

Очевидно, что приведенный рисунок дерева зависимостей, не отражает структурно-грамматических характеристик предложений: предложение P1 является односоставным определенно-личным, в то же время предложение P2, двусоставное полное, и предложение P3, двусоставное неполное без подлежащего, имеют такую же схему.

Если исходить из цели разработки графического представления синтаксических конструкций: «видеть структурные показатели текста» [4, с.96], которые позволят охарактеризовать синтаксический «почерк» автора, выявить индивидуальные особенности авторского стиля, то становится очевидным невозможность выявления в некоторых случаях индивидуально-авторских стилевых параметров, получается как бы «графическая омонимия»: разные типы предложений имеют одинаковую модель. В функциональной стилистике накоплен большой опыт по параметризации стилей, при этом отмечается, что все существующие параметры имеют «априорно равное право на существование (с точки зрения лингвистической обоснованности) в системе описания стиля» [2, с.67], а параметры синтаксического уровня языка обладают высокими стиледифференцирующими возможностями. В связи с этим выбор синтаксических параметров и их описание является одним из самых значимых при анализе стиля, имеющим четкие формальные критерии отбора с определением всех переходных случаев. Предлагаемые уточнения позволят моделировать структурную схему простых предложений исходя из отражения в структуре деревьев синтаксических категорий словесных позиций, а не слов. Рассмотрим конструкции и некоторые соглашения, позволяющие по рисунку дерева определять грамматическую оформленность предикативного центра (двусоставное или односоставное), тип односоставного предложения, а также структурную полноту или неполноту любого простого предложения.



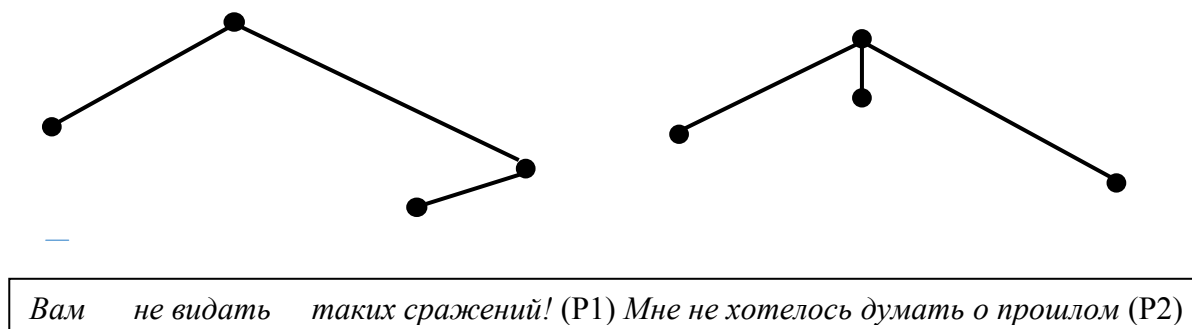
Известно, что полные односоставные предложения могут быть однословными. Разберем несколько примеров: *Встать!* (P1); *Темнеет* (P2); *Морозно* (P3); *Становится теплее* (P4). В первом предложении монопредикат – независимый инфинитив, выражающий значения императива, а три следующих предложения с точки зрения семантики и грамматики одинаковые: состояние окружающей среды, безличные, структура сформирована одной синтаксической позицией. Однако неодинаковая выраженность монопредиката: полнозначный безличный глагол – формально простое глагольное сказуемое, категория состояния – формально составное именное сказуемое с нулевой связкой и составное именное сказуемое, состоящее из глагола с ослабленным лексическим значением и категории состояния – должны отражаться в структуре графа по-разному, чтобы выступать в качестве считываемых стилистических параметров. Ранее для составных предикатов было введено идентифицирующее обозначение: от узла, обозначающего корень дерева, – грамматически главное слово в составном глагольном и именном предикате, по вертикали вниз на один уровень (на уровень подлежащего) находится узел, обозначающий зависимый компонент – инфинитив или именную часть [6, с.857 ]. (рис.2).



**Рис. 2. Структурные схемы односоставных нераспространенных (инфинитивного и безличных) предложений**

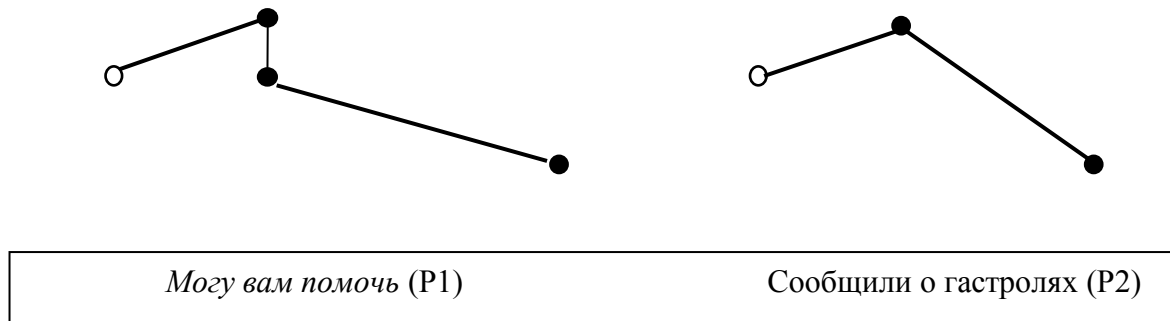
Графы односоставных безличных и инфинитивных предложений не содержат узла, соответствующего позиции подлежащего (корнем дерева является сказуемое, подлежащее обозначается узлом, всегда находящимся на уровень ниже сказуемого, независимо от его положения по отношению к сказуемому, ребро под углом), поэтому предложения P1, P2 и P3 представлены пустыми графами, P4 имеет одно ребро: вертикаль вниз на один уровень (на уровень подлежащего).

Распространенные односоставные предложения, в нашем примере инфинитивное и безличное: *Вам не видать таких сражений!* (P1), *Мне не хотелось думать о прошлом* (P2) - имеют следующий вид (рис.3).



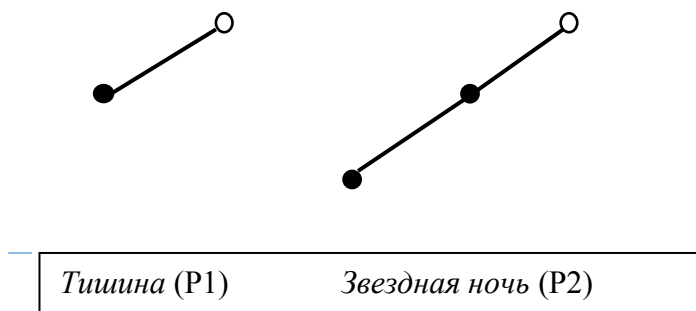
**Рис. 3. Структурная схема схемы односоставных распространенных (инфинитивного и безличного) предложений**

Односоставные определенно-личные и неопределенно-личные предложения: *Могут вам помочь* (P1). *Сообщили о гастроях* (P2) - в структуре графа будут содержать мнимый узел, соответствующий подлежащему (не закрашенный круг). Предложение P1 содержит сказуемое, разорванное дополнением; ребро соединяет лексически главный компонент сказуемого, от которого задается вопрос к дополнению, и дополнение (узлы, соответствующие синтаксическим категориям второстепенных членов, всегда находятся ниже уровня подлежащего) (рис.4).



**Рис. 4. Структурная схема схемы односоставных распространенных (определенно-личного и неопределенно-личного) предложений**

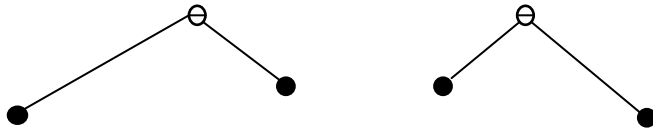
В графы, изображающие односоставные номинативные предложения, также предлагается ввести мнимый узел (не закрашенный круг), соответствующий глаголу-сказуемому *есть*: *Солнце* (P1). *Звездная ночь* (P2) (рис.5). Это необходимо для сохранения структуры дерева, которая подразумевает обязательное наличие вершины.



**Рис. 5. Структурные схемы односоставных номинативных (нераспространенного и распространенных) предложений**

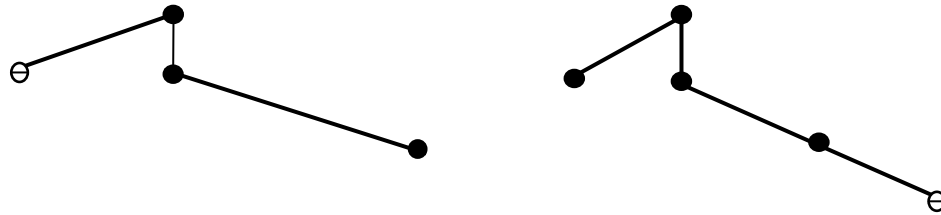
Введение в графы условных обозначений, позволяющих дифференцировать типы односоставных предложений, влечет за собой необходимость отражения в моделях неполноты структуры, кроме того, в качестве функционально-стилистической или авторской характеристики синтаксического «почерка» может быть введен такой параметр, как количество неполных синтаксических конструкций.

Предлагается ввести в графы простых неполных предложений, и двусоставных, и односоставных, дополнительный узел (узел с минусом), обозначающий отсутствующие члены предложения, создающие неполноту, в которые входят в сильные присловные распространители: прямое дополнение, обстоятельство при сказуемом, обозначающим движение в каком-то направлении, компонент несвободного словосочетания и др. На рис.6 представлены неполные предложения: *На небе звезды* (P1), *Я – в город* (P2), в структуре которых отсутствует сказуемое, и на рис.7 деревья зависимостей двусоставных предложений: *Хотел помочь вам* (P3), *Я должен все объяснить* (P4), в структуре которых соответственно отсутствуют подлежащее и дополнение.



*На небе звезды (P1) Я – в город (P2)*

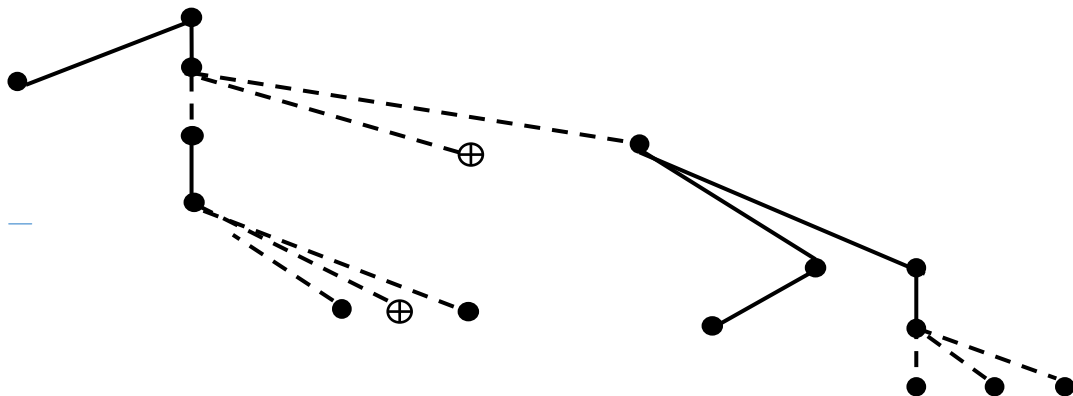
**Рис. 6. Структурные схемы неполных (без сказуемого) предложений**



*Хотел помочь вам (P3), Я должен все объяснить (P4)*

**Рис. 7. Структурные схемы неполных (без подлежащего и дополнения) предложений**

Однородность обозначается при помощи введения мнимого узла, ребра от которого изображаются пунктиром, узел, обозначаемый плюсом, соответствует сочинительным союзам, что соответствует условным обозначениям, предложенным И.П.Севбо, с небольшим уточнением: первый компонент ряда однородных членов соединяется пунктирным ребром с мнимым узлом по вертикали, на один уровень ниже (Рис.8).



*Солнце светило ласково и нежно и согревало своим теплом лес, луг, озеро*

**Рис. 8. Структурная схема простого предложения, осложненного однородными членами**

Полученное дерево зависимостей отражает три группы однородных членов в предложении, кроме того, можно с достаточной вероятностью видеть их синтаксическую позицию: однородные сказуемые, на что указывает мнимая вершина, идущая от корня дерева, и

однородные обстоятельства или дополнения, что можно предположить, исходя из анализа отраженных в схеме связей.

### **Выводы**

Таким образом, мы видим, что, если считать единицами таксономического уровня в дереве зависимостей не слова, а синтаксические позиции, что абсолютно соответствует организации этого языкового уровня, и принять некоторые соглашения, касающиеся изображения составных предикатов, синтаксически несвободных словосочетаний, занимающим позиции второстепенных членов предложения, отсутствующего второго главного члена в односоставных предложениях, синтаксических позиций, создающих неполноту структуры простого предложения, и ввести соответствующие условные обозначения, графы, репрезентирующие разные виды предложений, достаточно легко дифференцируются. Деревья зависимостей отражают не только специфику односоставных предложений в сравнении с двусоставными, но и их структурные особенности: тип предиката и монопредиката, характер неполноты, осложнение однородными членами, т.е. параметры, которые должны рассматриваться при описании авторского и функционального стиля и которые могут оказаться дифференцирующими. Предложения, имеющие осложнения структуры (обособленные определения, обстоятельства, дополнения, вводные, вставные конструкции) предполагается рассмотреть в дальнейшем.

### **Библиографический список**

1. **Бабайцева, В.В.** Синтаксис русского языка. М.: Флинта, 2015. – 576 с.
2. **Марусенко, М.А.** Атрибуция анонимных и псевдонимных произведений методами теории распознавания образов. Ленинград: ЛГУ, 1990. - 127
3. Русская грамматика. Синтаксис. Т.2. М.: Наука, 1980. – 709 с.
4. **Севбо, И.П.** Графическое представление синтаксических структур и стилистическая диагностика. Киев: Наукова думка, 1981. – 192 с.
5. **Синелева, А.В.** Атрибуция «Романа с кокаином»: лингвостатистическое исследование. Н.Новгород: ННГУ им.Н.И.Лобачевского, 2000. – 91с.
6. **Синелева, А.В.** Графическое моделирование синтаксических конструкций (простое предложение) // Информационные системы и технологии. 2020. С.854-862.
7. **Синелева, А.В.** Структура заголовочного комплекса в новостном дискурсе СМИ // Научный диалог. № 6. 2019. С. 72-86.

**СИСТЕМА ОБНАРУЖЕНИЯ И РАСПОЗНАВАНИЯ ДОРОЖНЫХ ОБЪЕКТОВ  
ПО МАРШРУТУ ДВИЖЕНИЯ ТРАНСПОРТНЫХ СРЕДСТВ**

А.А.Коротышева

ORCID: 0000-0001-8877-4235 e-mail: ania.korotishewa@yandex.ru

С.Н. Жуков

ORCID: 0000-0002-4410-4720 e-mail: jsn@rf.unn.ru

Нижегородский государственный университет им. Н.И. Лобачевского

Доклад посвящен созданию системы обнаружения и распознавания дорожной ситуации, основанной на обработке входящего потока видеoinформации от встроенных в транспортные средства видеокамер. Проведено исследование современных подходов к обработке изображений и построения моделей обнаружения и распознавания образов применительно к использованию на автомобильном транспорте. Выполнен обзор существующих наборов данных (датасетов), создан датасет с реальной дорожной обстановкой г. Нижнего Новгорода, на основе которого сформированы наборы для обучения, валидации и тестирования нейросетевой модели интеллектуальной системы. Разработаны процедуры для определения объектов, встречающихся на дороге (светофоры, дорожные знаки и т.д.), а также требования к архитектуре системы распознавания дорожной ситуации. Проведено тестирование интеллектуальной системы и выполнен анализ полученных результатов. Определены перспективы дальнейшей работы над улучшением показателей системы.

**Ключевые слова:** информационная система, набор данных, распознавание, нейронная сеть, дорожные знаки.

**THE SYSTEM OF DETECTION AND RECOGNITION OF ROAD OBJECTS ALONG THE  
ROUTE**

A.A. Korotysheva, S.N. Zhukov

Nizhniy Novgorod state university n. a. N.I. Lobachevsky

**Purpose:** The work is devoted to the creation of a system for detecting and recognizing a traffic situation based on processing the incoming video information stream from video cameras built into vehicles.

**Design / methodology / approach:** A study of modern approaches to image processing and building models for the detection and recognition of images in relation to use in road transport has been carried out.

**Findings:** A review of existing datasets was made, a dataset was created with the real traffic situation in Nizhny Novgorod, on the basis of which sets were formed for training, validating and testing a neural network model of an intelligent system. Procedures have been developed for determining objects encountered on the road (traffic lights, traffic signs, etc.), as well as requirements for the architecture of the traffic situation recognition system. The intelligent system was tested and the results were analyzed. The prospects for further work on improving the performance of the system are determined.

**Research limitations/implications:** In the future, the system will be modified to develop procedures for identifying road scenes in video information.

**Originality/value:** An original dataset was created based on the real traffic situation in Nizhny Novgorod for training, validation and testing of the neural network. The intelligent traffic situation recognition system contains original procedures for detecting objects encountered on the road.

**Key words:** information system, dataset, recognition, neural network, road signs.

## Введение

Обеспечение повышения безопасности дорожного движения в последнее время достигается за счет создания интеллектуальных навигационных систем, выполняющих анализ дорожной обстановки с использованием технологии искусственного интеллекта (англ. artificial intelligence, AI) [1]. Для решения задачи распознавания образов – определения объектов в потоке данных – стало популярным использование нейронных сетей, которые позволяют анализировать большие объемы данных. При этом на обнаружение объектов направлены модели-детекторы [2], а на определение классов объектов – классификаторы [3].

## Выбор модели распознавания

Одной из наиболее распространенных архитектур нейронных сетей на сегодняшний день, применяемых для распознавания объектов, является сеть YOLO (англ. You Only Look Once), разработанная Джозефом Редмонем и Али Фархади в 2016 году [4]. Модель сверточной нейронной сети YOLOv5 – усовершенствованная пятая версия YOLO, реализованная на фреймворке PyTorch [5].

Сеть может выдавать предсказания со скоростью 140 кадров в секунду, что обеспечивает возможность использования модели для получения предсказаний в режиме реального времени. Такая скорость работы достигается за счет архитектуры сети, которая позволяет получать результат по изображению за один проход через нейронную сеть, тогда как обычно изображение требуется анализировать дважды из-за последовательного применения детектора для определения областей, в которых может быть объект, и классификатора для определения класса объекта.

## Создание датасета для обучения нейросети

В России знаки дорожного движения соответствуют международным нормам, принятым на Венской конвенции. То есть они стандартизированы и не требуют знания национального языка. Но знаки некоторых стран, например, США, хоть и соответствуют стандартам, содержат другие цвета, большое количество слов и т.п. При этом светофоры являются более универсальным устройством регулирования дорожного движения, что позволяет использовать открытые датасеты со светофорами других стран.

Таким образом, для решения задачи одновременного обнаружения и распознавания, как светофоров, так и знаков дорожного движения в видеопотоке, был сформирован набор данных, совмещающий выборки Russian Traffic Signs (RTS) [6], часть German Traffic Signs (GTS) [7], другие знаки дорожного движения, встречающиеся на российских дорогах, а также значения сигналов светофоров, часть которых взята из датасета Bosch TrafficLights (BTL)[8].

Все кадры были приведены к единому разрешению 640x640 пикселей и аннотированы по правилам, принятым для обучения моделей YOLO. Аннотацией каждого изображения является текстовый файл формата \*.txt, название которого соответствует описываемому изображению. Файл включает описание объектов выбранных классов в следующем формате:

$$n, x, y, w, h.$$

Здесь  $n$  – номер класса объекта;  $x$  – относительная координата центра рамки, выделяющей объект, по оси  $x$ ;  $y$  – относительная координата центра рамки, выделяющей объект, по оси  $y$ ;  $w$ ,  $h$  – относительная ширина и высота рамки, выделяющей объект.

Для увеличения выборки часть изображений была аугментирована (приумножена, англ. augmentation), для этого к исходным изображениям применялись следующие функции:

- случайное вращение от -5 до +5 градусов;
- случайная регулировка яркости от -20 до +20 процентов;
- случайное размытие по Гауссу от 0 до 1 пикселя;
- к 2 процентам пикселей - шум соль и перец (случайно возникающие чёрные и белые пиксели);
- произвольная обрезка от 0 до 50 процентов изображения.

К выделяющим рамкам на изображениях были применены следующие преобразования:

- случайное вращение от -10 до +10 градусов;
- случайный сдвиг от  $-10^\circ$  до  $+10^\circ$  по горизонтали и по вертикали от  $-10^\circ$  до  $+10^\circ$ .

### Обучение нейросетевой модели

В качестве метрик оценки качества модели распознавания использовались стандартная метрика точности (англ. precision) и mAP@ [.5:.95] (англ. meanaverage precision), используемая при решении задачи классификации, если количество классов объектов больше двух – средняя по классам средняя точность, которая является результатом суммирования значений, полученных по метрике средней точности (англ. average precision, AP), и последующего деления суммы на количество всех классов [9].

Графики изменения метрик (precision, mAP@[.5:.95]) в процессе обучения на выбранных изображениях, представлены на рис.1, 2.

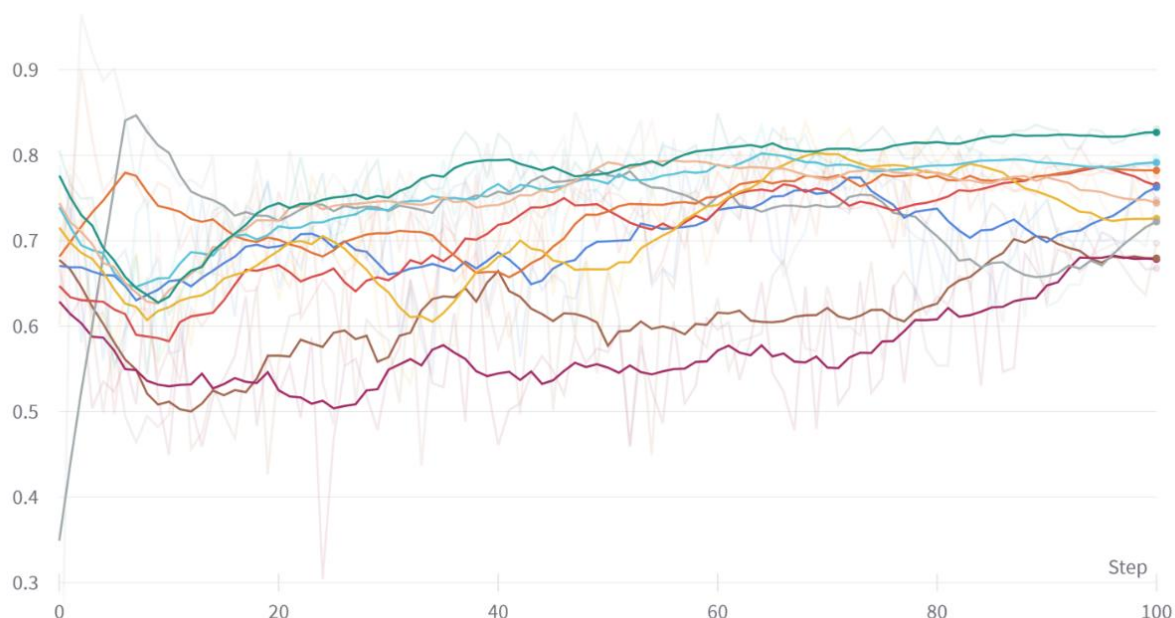
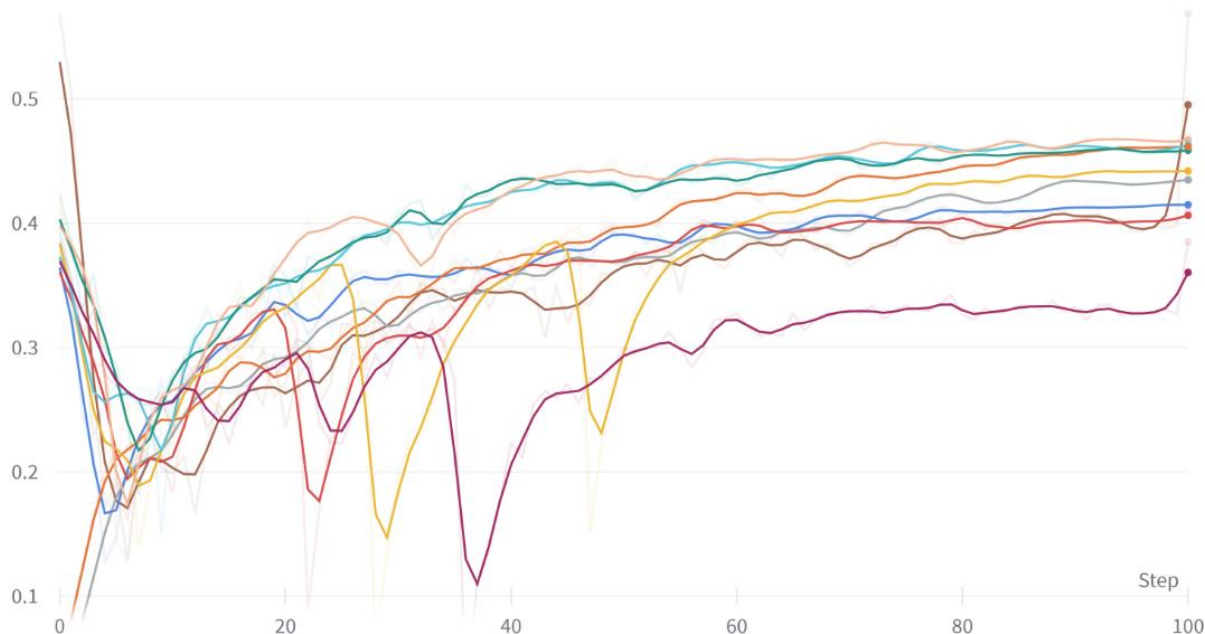


Рис.1. График изменения метрики точности



**Рис.2. График изменения метрики mAP@[.5:.95]**

### Оценка результатов тестирования

Модель должна обнаруживать и распознавать широкий спектр классов целевых объектов, включая знаки зоны действия, без которых понимание дорожной ситуации будет неполным, а в отдельных случаях в принципе невозможным.

Существующие датасеты содержат большое количество изображений, но не включают в разметку требуемые классы дорожных знаков (табл. 1).

**Таблица 1**

**Сравнение наборов данных. \* - зависит от выборки**

Название датасета	Изображений	Подходит
German TS	900	Нет
Russian TS	179138	Нет
Chinese TS	110000	Нет
LISATS [10]	6610	Нет
Сформированный набор данных	*	Да

Рассмотрим результаты распознавания моделей, обученных на наборах German Traffic Signs, Russian Traffic Signs и сформированном наборе данных на следующем кадре реальной дорожной обстановки (рис. 3).





**Рис.3 – Пример кадра реальной дорожной обстановки**

Все параметры, задаваемые при обучении моделей, кроме обучающего набора (он включен в конфигурационный файл) инициализированы одинаково, основные параметры приведены в табл. 2.




**Таблица 2**

**Параметры, заданные при обучении сети YOLOv5**

Описание	Имя параметра	Значение
Размерность входного тензора – ширина и высота изображений на входе в нейронную сеть, количество каналов в изображении	img	(640, 640, 3)
Размер подвыборки	batch	16
Веса предобученной модели	weights	imagenet
Функция ошибки	loss	categorical_crossentropy
Оптимизатор - тип оптимизатора	optimizer	Adam
Скорость обучения оптимизатора	lr	1e-5
Количество эпох	epochs	300
Путь до конфигурационного файла	data	dataset.yaml

Модели на приведенном примере выдают следующие результаты (табл. 3).

Сравнение результатов работы моделей

Датасет	Результат распознавания
GTS	
RTS	
Сформированный набор данных	

По данным из таблицы видно, что результатом работы первой модели на первом (верхнем) объекте является класс, обозначающий другой знак дорожного движения, второй объект (средний) не был определен как знак, а класс третьего объекта (нижнего) – не задан в разметке этого набора данных. Результат второй модели: детектированы два первых объекта с максимальными вероятностями (больше 0,8) для классов: 4\_1\_1 или «Движение прямо» и 3\_27 – «Остановка запрещена» соответственно, что совпадает с реальными классами объектов. Третий результат – все три знака детектированы, классы совпадают с реальными и их вероятности больше 0,9.

Таким образом, модель, обученная на сформированном наборе данных, является более точной и полной (охватывает большее количество классов объектов) по сравнению с моделями, обученными на известных датасетах.

### Заключение

Система обнаружения и распознавания дорожных объектов по маршруту движения транспортных средств повысит безопасность перемещения беспилотных автомобилей по дорогам за счет точного обнаружения дорожных объектов в режиме реального времени.

В дальнейшем планируется расширение набора данных за счет большего числа дорожных объектов, встречающихся по маршруту движения транспортных средств, а также увеличение разнообразия погодных условий, при которых система будет работать с достаточной точностью.

### Библиографический список

1. **Hamdan, M., Shehadeh, H.** Automated Quantification of Eye Blink Rate Using VIOLA-JONES Algorithm // International Journal of Technology Diffusion. 2018. № 4. С. 19-25.
2. **Лашков, И.Б.** Детектирование состояния зевоты у водителя транспортного средства при помощи модели сверточной нейронной сети // Научно-технический вестник информационных технологий, механики и оптики, vol. 22, no. 1, 2022. P. 33-46.
3. **Zolotenkova, G.V., Rogachev, A.I., Pigolkin, Y.I., Edelev, I.S., Borshchevskaya, V.N., Cameriere, R.** Age Classification in Forensic Medicine Using Machine Learning Techniques // Современныетехнологии в медицине, vol. 14, no. 1 (eng), 2022. P. 15-22.
4. **Redmon, J., Farhadi, A.** YOLO9000: Better, Faster, Stronger // IEEE Conference on Computer Vision and Pattern Recognition, 2016.
5. **Redmon, J., Divvala, S., Girshick, R., Farhadi, A.** You Only Look Once: Unified, Real-Time Object Detection // Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2016.
6. **Shakhuro, V.I., Konushin, A.** Russian traffic sign images dataset // Computer Optics, 40, 2016. P.294-300.
7. **Houben, S., Stallkamp, J., Salmen, J., Schlipsing, M., Igel, C.** The German Traffic Sign Recognition Benchmark: A multi-class classification competition // Conference: Neural Networks (IJCNN), 2011.
8. **Karsten, B., Novak, L., Botros, R.** A deep learning approach to traffic lights: Detection, tracking, and classification // 2017 IEEE International Conference on Robotics and Automation (ICRA), 2017. P. 1370-1377, doi: 10.1109/ICRA.2017.7989163.
9. **Everingham, M., Gool, L. V. , Williams, C. K. I., Winn, J., Zisserman, A.** The PASCAL Visual Object Classes (VOC) Challenge // International Journal of Computer Vision, 88, 2010. P. 303–338.
10. **Mogelmoose, A., Trivedi, M. M., Moeslund, T.B.** Vision based Traffic Sign Detection and Analysis for Intelligent Driver Assistance Systems: Perspectives and Survey // IEEE Transactions on Intelligent Transportation Systems, 2012.

**ИСПОЛЬЗОВАНИЕ ГЕНЕРАТИВНО-СОСТЯЗАТЕЛЬНЫХ НЕЙРОННЫХ СЕТЕЙ КАК СПОСОБ ПОВЫШЕНИЯ УСТОЙЧИВОСТИ РАСПОЗНАВАНИЯ ИЗОБРАЖЕНИЯ****А.А. Тищенко**

ORCID: 0000-0002-3236-0874 e-mail: and.tis@mail.ru

**В.Е. Гай**

ORCID: 0000-0002-3644-5234 e-mail: iamuser@inbox.ru

**Р.О. Баринов**

ORCID: 0000-0002-0269-3205 e-mail: barinovr@list.ru

Нижегородский государственный технический университет им. П. Е. Алексеева

Интеграция нейронных сетей для решения задач распознавания объектов все носит массовый характер вне зависимости от класса объектов. Такой рост их использования так или иначе создаст класс методов, целью которых является нарушение их работы. Примером таких методов является, к примеру состязательные атаки, атаки одним пикселем и прочие. Путем подачи на вход сверточной нейронной сети атакующего изображения нарушается работа сети и как следствие всей системы, основанной или включающей в себя нейронную сеть. В данном докладе будет дано определение устойчивости сверточной нейронной сети как параметра, рассмотрены различные способы противодействия атакам в том числе и с помощью генеративно-состязательных нейронных сетей.

**Ключевые слова:** устойчивость, сверточные нейронные сети, состязательные примеры

**GENERATIVE-ADVERSARIAL NEURAL NETWORKS TO IMPROVE THE STABILITY OF IMAGE RECOGNITION****A.A.Tischenko, V.E. Gai, R.O. Barinov**

Nizhniy Novgorodstate technical university n.a. R.E.Alekseev

**Purpose:** This article discusses the use of generative-adversarial neural networks to improve the stability of convolutional neural networks by creating/increasing a data set with adversarial examples. We propose the concept of a system that is able to recognize images even after adversarial attacks on the neural network.

**Design/methodology/approach:** The article will analyze several ways of integrating a generative-adversarial neural network into a cascade with a convolutional neural network at different stages of image recognition: from training the network to its practical application.

**Findings:** A new image recognition system architecture with the ability to integrate actor-based resilience techniques is proposed. Different ways of integration are considered: before training - image processing, during training - using a dataset created with GAN's, during use - the image is processed before entering the input of the convolutional neural network.

**Research limitations/implications:** The present article will provide a basis for the development of software or systems to improve the stability of convolutional neural networks in solving practical problems

**Originality/value:** The system based on the considered methods can be applied in a real-time face recognition system or in a pass control system, based on neural networks.

**Keywords:** stability, convolutional neural networks, adversarial examples

## **Введение**

Рост интеграции нейросетевых технологий в массово используемые продукты, приложения и системы с каждым годом бьет рекорды. Как и любой распространенный тип ПО, системы, работа которых основана на применении нейросетевых алгоритмов, могут быть подвержены внешним воздействиям со стороны злоумышленников. Примером таких воздействий для сверточных нейронных сетей служат соревновательные атаки – изображение, поверх которого накладывается «шум», в виде изменения пикселя(-ей). Из-за данного зашумления происходит ошибка классификации. Поэтому в ряде систем, ключевым элементом, которых является классификация и как следствие идентификация некоторой уникальной информации (лицо, отпечаток пальцев, голос, QR-код и т.д.), наравне со стандартным для нейросети параметром – точности, необходим ввод такого параметра как устойчивость нейронной сети.

Генеративно-состязательные нейронные сети могут внести существенный вклад в повышение устойчивости – благодаря большому количеству данных повышается абстрактность мышления нейронной сети, и как следствие устойчивость к разного рода соревновательным атакам (как естественным, так и описанным выше генерируемым шумам).

## **Подготовка к эксперименту**

Для оценки эффективности использования генеративно-состязательных нейронных сетей в повышении устойчивости необходимо выполнить следующие шаги:

1. Выбор сверточной нейронной сети с готовой архитектурой.
2. Введение новых метрик для отслеживания эффективности повышения устойчивости.
3. Выбор генеративно-состязательной нейронной сети.

Общий ход работы будет происходить следующим образом: выбранная нейронная сеть обучается на трех разных наборах данных (стандартный, с дополнением сгенерированных данных, с добавлением данных, созданных стандартными средствами аугментации), использовать методы генерации соревновательных атак и обработать моделями новые данные с применением введенных метрик, полученный результат сравнить и проанализировать.

Для разработки экспериментальной части использовалась библиотека Keras на языке программирования Python. Данная библиотека подходит для создания прототипов за счет удобства использования готовых архитектур нейросетей.

В качестве тестируемой сверточной нейронной сети была выбрана модель MobileNet [1]. Данная модель легковесна и быстро обучается, за счет скорости обработки данных её основное применение ограничивается применением в сфере мобильных приложений, которые в свою очередь являются наиболее используемой технологией применения нейронных сетей для обычных пользователей (сканер отпечатков пальцев, разблокировка смартфона с помощью распознавания лица и пр.).

Для генерации новых данных использовалась генеративно-состязательная нейронная сеть StarGANv2. Ее принцип работы и особенности будут описаны далее.

В качестве исходного набора данных использовалось классическое соревнование KaggleCatvs. Dog. Использование набора данных с большим количеством классов будет неэффективно, так как время обучения и классификации в ходе эксперимента возрастут, а необходимость есть только в отслеживании «попадания» в наш класс – количество неверных классов значения не имеют.



В качестве методов генерации соревновательных примеров будут использованы метод быстрого градиентного спуска и его модификации: FGM, FGSM, I-FGSM(PGD)[2], а также One-PixelAttack. Краткое описание этих методов будет представлено далее.

Новые метрики представляют собой модификацию уже существующей классической метрики для решения задачи классификации *accuracy* и метрику, основанную на разнице вероятностей.

### Модели

MobileNet [1] -сверточная нейронная сеть, предназначенная для классификации изображений в мобильных приложениях. Быстродействие ей обеспечивает отсутствие слоев подвыборки, их заменяет слои свертки со параметром снижения пространственной размерности. Настраивается с помощью двух параметров  $\alpha$  и  $\rho$ , множитель ширины и глубины соответственно. Множитель ширины отвечает за количество каналов в каждом слое, множитель глубины за пространственные размеры входных тензоров. Увеличение этих параметров способно существенно повысить точность распознавания, но в то же время снижает скорость работы. В рамках эксперимента использовались рекомендуемые параметры:  $\alpha = 1$ ,  $\rho = 1$ . Архитектура представляет собой чередование 13 блоков: слой глубокой связанной свертки, слой нормализации, функцией активации ReLU, слой обычной свертки, слой нормализации и функция активации. Схема архитектуры изображена на рисунке 1[1]. На выходном слое располагаются два нейрона (в соответствии количеству классов) и функция активации softmax, возвращающая массив вероятностей принадлежности каждому классу.

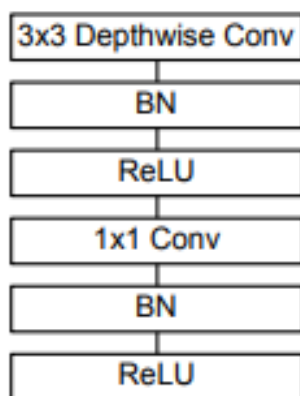


Рис. 1. Архитектура MobileNet

Модель StarGANv2 [3] – генеративно-сопоставительная нейронная сеть, принцип работы заключается в «совмещении» признаков двух поданных на вход изображений. Первое изображение называется основой, второе вдохновителем. Сверточная нейронная сеть, входящая в каскад генеративно-сопоставительной сети, выделяет основные признаки из изображения-вдохновителя, а затем генератор «накладывает» их на изображение-основу. Новое изображение подается на вход сверточной нейронной сети, которое классифицирует объект на изображении. В случае если класс нового и старого изображения совпадает, то новое изображение повторяет описанные выше шаги, в противном случае генерация изображения считается завершенным и поступает на выход каскада. В рамках эксперимента изображение-основа и изображение-вдохновитель должны принадлежать одному классу для формирования корректного дополненного набора данных.

### Методы генерации соревновательных примеров

Методы FGSM, I-FGSM(PGD), MI-FGSM [2], являются основными методами модификации изображения, принцип из работы основан на вычислении градиента функции ошибки. У входного изображения вычисляется карта градиентов функции ошибки, которая затем на него накладывается, тем самым получается модифицированное изображение. Степень модификации зависит от параметра  $\varepsilon$ , регулирующий интенсивность накладываемого «шума». Так как изменения в изображении должны быть незаметными для человеческого глаза в рамках эксперимента данный параметр  $\varepsilon = 0.001$ .

1. Метод FGSM использует линейную часть функции потерь в входном изображении и вычисляется по следующей формуле:

$$X' = X + \varepsilon * \text{sign}(\nabla_x J(X, y_{true}))$$

где  $X'$  – модифицированное изображение,  $X$  – исходное изображение,  $\varepsilon$  – коэффициент интенсивности (от 0 до 1),  $y_{true}$  – истинная метка класса.

2. Метод I-FGSM – итеративная вариация метода FGSM, выражается следующей формулой:

$$X^{n+1} = \Pi_B(X^n + \varepsilon * \text{sign}(\nabla_x J(X, y_{true})))$$

где  $n$  – номер итерации от 0 до  $N$ ,  $N$  – количество итераций.

3. Метод MI-FGSM основан на I-FGSM, со сглаженным градиентом, вышеописанная формула преобразуется следующим образом:

$$X^{n+1} = \Pi_B(X^n + \varepsilon * \text{sign}(g^{t+1}))$$

$$g^{n+1} = \mu * g^n + \frac{\nabla_x J(X, y_{true})}{\|\nabla_x J(X, y_{true})\|}$$

где  $g$  – сглаженный градиент,  $\mu$  – коэффициент сглаживания.

Данные методы реализованы в библиотеке FoolBox.

### Метрики

Метрика *accuracy*[4] используется для отслеживания точности классификации нейронной сети и описывается следующим соотношением и матрицей ошибок.

$$\text{accuracy} = \frac{TP + TN}{TP + TN + FP + FN}$$

где  $TP$ ,  $TN$ ,  $FP$ ,  $FN$  – результаты классификации, представленные в матрице ошибок (Таблица 1)

Таблица 1.

Матрица ошибок

	$Y = 1$	$Y = 0$
$\hat{Y} = 1$	True Positive (TP)	False Positive (FP)
$\hat{Y} = 0$	False Negative (FN)	True Negative (TN)

В таблице  $Y$  – результат классификации, который должен быть.  $\hat{Y}$  – результат классификации, полученный на выходе нейронной сети. На основе метрики *accuracy*, предлагаемая метрика устойчивости *adv\_accuracy* будет другое описание матрицы ошибок: обозначим  $Y$  – результат классификации объекта, а  $\hat{Y}$  – результат классификации того же самого объекта, но атакованного, некоторым алгоритмов генерации шума. В отличие от оригинальной метрики *accuracy*, описанная выше метрика *adv\_accuracy* не имеет недостаток

в виде требования равных классов, так как матрица ошибок строго бинарная – определяется факт совпадения с исходным классом. Рассматриваемый параметр устойчивости нейронной сети прямо пропорционален значению введенной метрики.

Вторая метрика больше подойдет к моделям регрессии, но может быть применена и для сверточной нейронной сети. Представляет собой среднее арифметическое разницы вероятностей классификации стандартного и соревновательного объекта в пределах одного класса и может быть выражена формулой.

$$adv\_diff = \frac{\sum_{i=0}^n (P_i - P_{adv_i})}{n}$$

где  $P_i$  – вероятность классификации исходного изображения,  $P_{adv_i}$  – вероятность классификации соревновательного примера,  $n$  – размер набора данных (количество обработанных изображений).

### Результаты экспериментов

В ходе экспериментов были обучены модели Mobile Net на трех разных наборах данных (стандартный набор, с применением генеративно-сопоставительной нейросети, с применением стандартных методов аугментации), результаты обучения представлены на рис. 2.

```
Epoch 50/50
586/586 [=====] - 1371s 2s/step - loss: 0.0148 - accuracy: 0.9946 - val_loss: 0.2679 - val_accuracy:
0.9389
: <keras.callbacks.History at 0x205d7640dc0>

Epoch 50/50
844/844 [=====] - 2195s 3s/step - loss: 0.0059 - accuracy: 0.9978 - val_loss: 0.2032 - val_accurac
y: 0.9587
<keras.callbacks.History at 0x2071b14fa30>

Epoch 50/50
586/586 [=====] - 1435s 2s/step - loss: 0.0429 - accuracy: 0.9830 - val_loss: 0.0865 - val_accurac
y: 0.9690
<keras.callbacks.History at 0x2071a32f5b0>
```

**Рис. 2. Результаты обучения моделей. сверху: стандартный набор данных, посередине: дополненный модифицированными изображения, снизу: с применением стандартных методов аугментации.**

**Таблица 2.**

### Результаты экспериментов с введенными метриками.

	acc	FGSM		I-FGSM		MI-FGSM	
		adv_acc	adv_diff	adv_acc	adv_diff	adv_acc	adv_diff
MobileNet	0.9389	0.3423	0.6106	0.2716	0.6823	0.2301	0.7742
MobileNet_GAN	0.9587	0.5398	0.4431	0.4928	0.5006	0.3920	0.5927
MobileNet_stAug	0.9690	0.4035	0.5829	0.3124	0.6529	0.2836	0.7342

Результаты обучения и использования новых метрик для каждой модели были сведены в следующую таблицу (табл.2). Где adv\_acc – новая метрика, основанная на метрике accuracy (в таблице указана как acc), adv\_diff – новая метрика основанная на среднем



арифметическом разности вероятностей классификации. Результаты указаны для всех трех рассмотренных ранее методов генерации соревновательных примеров.

### **Выводы**

На основе полученных результатов можно сделать вывод, что использование генеративно-сопоставительных нейронных сетей потенциально способно эффективно повышать устойчивость нейронных сетей перед соревновательными атаками. Если сравнить результаты применения генеративно-сопоставительных нейросети и стандартных метод аугментации, то можно увидеть, что по всем параметрам использование стандартных метод аугментации менее эффективно, однако применение генеративно-сопоставительных нейронных сетей затрачивает больше ресурсов и времени. Примененные методы расширения набора данных показали свою эффективность, однако, в случае необдуманного использования они могут привести к полной неработоспособности нейронной сети, так как чем больше объектов приходится на каждый класс, тем выше вероятность переобучения.

### **Библиографический список**

- 1. G. Howard, Andrew, Zhu, Menglong, Chen, Bo, Kalenichenko, Dmitry, Wang, Weijun, Weyand, Tobias, Andreetto, Marco, Hartwig Adam.** Mobile Nets: Efficient Convolutional Neural Networks for Mobile Vision Applications. - 17.04.2017 – Vol. arXiv: 1704.04861 – version 1. Access mode: <https://arxiv.org/pdf/1704.04861.pdf>.
- 2. Dong, Yinpeng, Liao, Fangzhou, Pang, Tianyu, Su, Hang, Zhu, Hu, Jun Xiaolin , Li., Jianguo** Discovering Adversarial Examples with Momentum – 17.10.2017 – Vol. arXiv: 1710.06081 – version 1. Access mode: <https://arxiv.org/pdf/1710.06081v1.pdf>.
- 3. Choi, Yunjey, Uh, Youngjung, Yoo, Jaejun, Ha, Jung-Woo.** StarGAN v2: Diverse Image Synthesis for Multiple Domains – 26.04.2020 – Vol. arXiv: 1912.01865 – version 2. Access mode: <https://arxiv.org/pdf/1912.01865.pdf>.
- 4. Grandini, Margherita, Bagli, Enrico, Visani, Giorgio.** METRICS FOR MULTI-CLASS CLASSIFICATION: AN OVERVIEW – 13.08.2020 -Vol.arXiv: 2008.05756 – version 1. Access mode: <https://arxiv.org/pdf/2008.05756.pdf>.

## СРАВНИТЕЛЬНЫЙ АНАЛИЗ ДИНАМИКИ РАСПРОСТРАНЕНИЯ ПЕРВОЙ ВОЛНЫ COVID-19 В ФЕДЕРАЛЬНЫХ ОКРУГАХ РОССИЙСКОЙ ФЕДЕРАЦИИ

М.В. Кокоулина<sup>1</sup>

ORCID: 0000-0001-5890-3649 e-mail: kokoulinamaria97@gmail.com

Е.Н. Лешехва<sup>1</sup>

ORCID: 0000-0003-2436-3431 e-mail: leshekhva.elena@mail.ru

Е.Н. Пелиновский<sup>1-3</sup>

ORCID: 0000-0002-5092-0302 e-mail: pelinovsky@gmail.com

А.С. Епифанова<sup>1</sup>

ORCID: 0000-0003-1061-0964 e-mail: epifanova.anastasia.s@gmail.com

А.А. Куркин<sup>1</sup>

ORCID: 0000-0003-3828-6406 e-mail: aakurkin@gmail.com

О.Е. Куркина<sup>1</sup>

ORCID: 0000-0002-4030-2906 e-mail: oksana.kurkina@mail.ru

<sup>1</sup>Нижегородский государственный технический университет им. Р. Е. Алексеева

<sup>2</sup>Национальный исследовательский университет «Высшая школа экономики»

<sup>3</sup>Институт прикладной физики РАН

В работе проведен сравнительный анализ динамики распространения первой волны COVID-19 в федеральных округах России. Аппроксимация роста числа заболевших выполнена с помощью простой и обобщенной логистических моделей, а также модели Гомпертца. Их параметры определяются регрессионным анализом на основе официальных данных Коммуникационного центра Правительства Российской Федерации, опубликованных на интернет-ресурсе стопкоронавирус.рф. Показано, что для всех округов РФ обобщенное логистическое уравнение является наилучшей регрессионной моделью из рассмотренных ( $R^2 > 0.94$ ). Модель Гомпертца имеет меньшее количество подбираемых параметров, однако ее коэффициент детерминации превышает 0.91.

**Ключевые слова:** математическое моделирование, простая логистическая модель, обобщенная логистическая модель, модель Гомпертца, COVID-19

## COMPARATIVE ANALYSIS OF THE COVID-19 FIRST WAVE SPREAD DYNAMICS IN THE FEDERAL DISTRICTS OF THE RUSSIAN FEDERATION

M.V. Kokoulina<sup>1</sup>, E.N. Leshekhva<sup>1</sup>, E.N. Pelinovsky<sup>1-3</sup>, A.S. Epifanova<sup>1</sup>, A.A. Kurkin<sup>1</sup>, O.E. Kurkina<sup>1</sup>

<sup>1</sup>Nizhny Novgorod State Technical University n.a. R.E. Alekseev, Nizhny Novgorod

<sup>2</sup>National Research University – Higher School of Economics, Nizhny Novgorod

<sup>3</sup>Institute of Applied Physics, Nizhny Novgorod

**Purpose:** This paper provides a comparative analysis of the COVID-19 first wave spread dynamics in the federal districts of Russia.

**Design/methodology/approach:** simple logistic, generalized logistic models and the Gompertz equations were used in this work. The initial data are taken from the official website стопкоронавирус.рф.

**Findings:** It is shown that for all considered districts the generalized logistic equation is a good regression model ( $R^2 > 0.94$ ).

**Research limitations/implications:** The present work provides a starting-point for further research of coronavirus infection waves in all regions of the Russian Federation.

**Originality/value:** The obtained results may be used for the modeling of COVID-19 waves in all regions of the Russian Federation.

**Keywords:** mathematical modeling, simple logistic model, generalized logistic model, Gompertz model, COVID-19

## Введение

Возникшая в конце 2019 г. в китайском городе Ухань эпидемия COVID-19, охватила уже буквально все страны мира и продолжается до сих пор. В связи с этим усилия большого количества специалистов направлены на изучение данной инфекции и на построение математических моделей, описывающих динамику роста заболеваемости и предсказание ее дальнейшего развития. Российская Федерация на данный момент занимает седьмое место по количеству заболевших и четвертое – по количеству умерших в мире. Ежедневно НИЛ «Моделирование природных и техногенных катастроф» НГТУ им. Р.Е. Алексеева занимается сбором данных и в качестве информационной базы исследования были использованы данные с сайта [Стопкоронавирус.рф](http://stopcoronavirus.rf). - официального интернет-ресурса о заболеваемости COVID-19 в Федеральных округах Российской Федерации. Накопленные за три года данные позволяют представить анализ в широком диапазоне значений от начала эпидемии до настоящего момента.

В настоящей работе на основе имеющегося на сегодняшний день массива данных по заболеваемости COVID-19 в регионах России произведен подбор наилучшей аппроксимации с помощью трех логистических моделей: классическая и обобщенная логистические модели и модель Гомпертца. Аналогичная работа была проведена ранее для двенадцати стран мира в [1,2, 3], а также в целом для России[4], и как показал анализ, для всех рассмотренных стран обобщенная логистическая кривая является неплохой аппроксимацией исходных данных. Стоит отметить, что анализ в рамках простых малопараметрических моделей важен, поскольку позволяет достаточно быстро и с удовлетворительной точностью описать процесс.

### Логистические модели распространения эпидемий

Для объяснения скорости распространения эпидемий и прогнозирования их последствий используется ряд математических моделей различной сложности. В данной работе мы остановились на рассмотрении трех простейших моделей.

Исторически первой моделью является простое логистическое уравнение [5], представляющая собой нелинейное ОДУ первого порядка с постоянными коэффициентами. Качественно она объясняет рост числа заболевших от времени: экспоненциальный рост числа заболевших на начальном этапе развития эпидемии и выход на константу к концу эпидемии. К удивлению многих, оказалось, что даже простейшая логистическая модель, неплохо описала численность переболевших в различных странах во время первой волны эпидемии [1, 6, 7]. Логистическое уравнение имеет вид

$$K = rN \left( 1 - \frac{N}{N_{\infty}} \right), \quad (1)$$

где  $K = dN/dt$  – количеством заболевших в один день,  $N$  – число переболевших людей,  $r$  – коэффициент роста и  $N_{\infty}$  - максимальное число переболевших.

В случае, когда на начальном этапе рост числа заболевших от времени описывается степенной зависимостью, используют обобщенное логистическое уравнение. В отличие от простой модели, обобщенное имеет большее число констант – параметров. Рассмотрим обобщенную модель логистического уравнения, содержащего четыре константы [8, 9]

$$K = rN^\alpha \left(1 - \frac{N}{N_\infty}\right)^\beta, \quad (2)$$

При  $\alpha = \beta = 1$  уравнение (2) совпадает с простым логистическим уравнением (1), решение которого тривиально находится [1].

В качестве третьей модели рассмотрено уравнение Гомпертца, предложенное им еще в 1825 году [10], оно использует те же две константы, что и простое логистическое уравнение. Однако оно не так активно используется в динамике развития эпидемий. Уравнение Гомпертца [10], записанное в стандартной форме, имеет вид:

$$K = rN \left(1 - \frac{\ln N}{\ln N_\infty}\right), \quad (3)$$

Отметим, что число констант здесь такое же как в классическом логистическом уравнении – их две.

Стандартный подход к определению коэффициентов модели состоит в минимизации среднеквадратического отклонения между модельными и реальными данными. Точность аппроксимации данных определяется с помощью коэффициента детерминации по формуле:

$$R^2 = 1 - \frac{\sum_{i=1}^n (K_i - \tilde{K}_i)^2}{\sum_{i=1}^n (K_i - \bar{K})^2}, \quad (4)$$

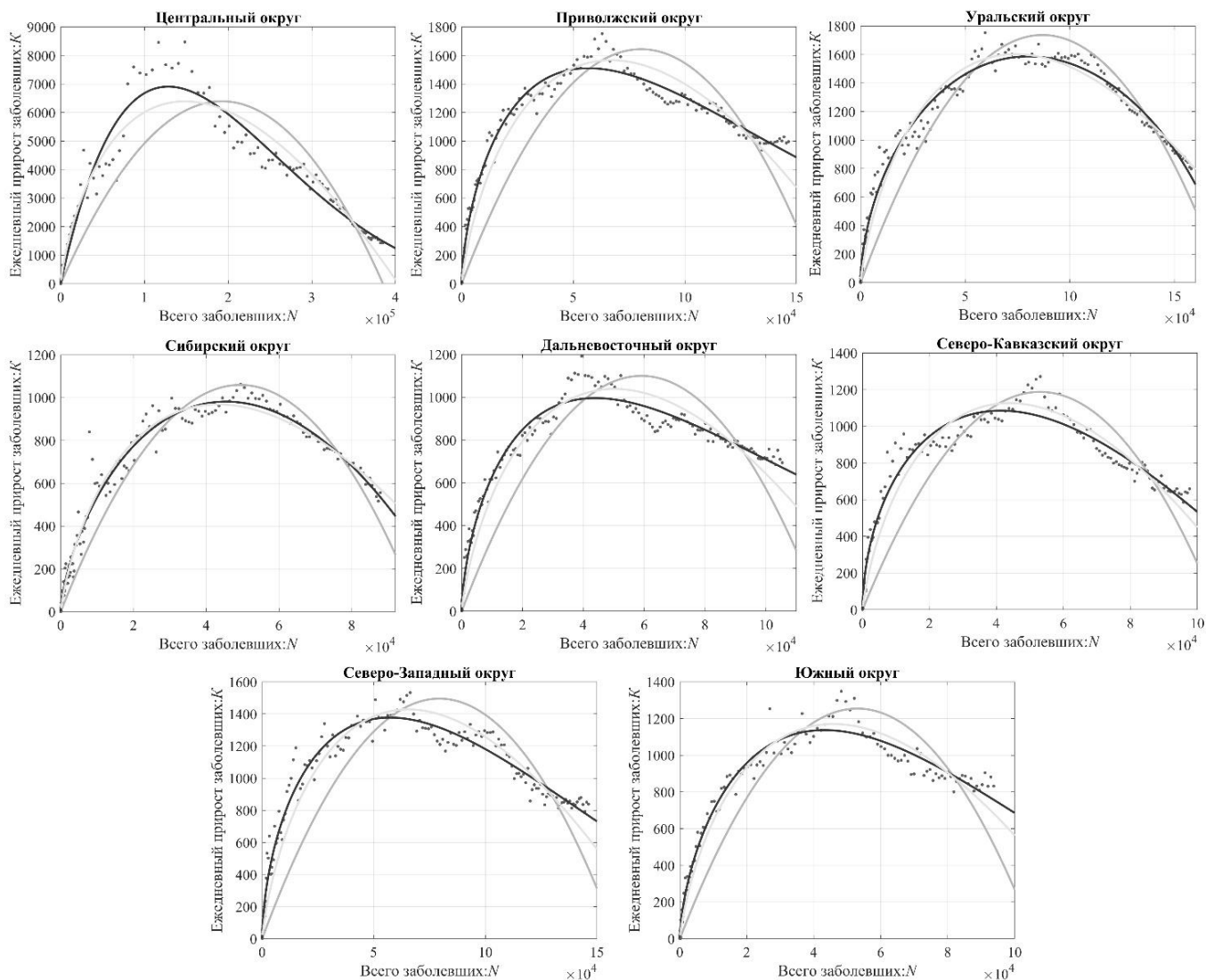
где  $K_i$  – количество заболевших в  $i$ -й день,  $\tilde{K}_i$  – модельное значение,  $\bar{K}$  – среднее значение заболевших за весь рассматриваемый период.

## Результаты

Рассмотрим применимость этих моделей для описания развития эпидемии COVID-19 на примере «первой волны» коронавирусной инфекции во всех округах Российской Федерации. На рис. 1 представлены данные медицинской статистики, которые отмечены синими точками, и всех трех алгебраических аппроксимаций для всех округов Российской Федерации. Здесь зеленая кривая соответствует классическому логистическому уравнению, красная – обобщенному логическому уравнению и желтая – уравнению Гомпертца. В табл. 1 приведены значения коэффициента детерминации  $R^2$ , который для логистической кривой (2) всюду превышает 0.9, что свидетельствует о достаточно хорошей аппроксимации. В то же время, если бы мы ограничивались простой логистической кривой, то коэффициент  $R^2$  в ряде случаев опускался бы до 0.7. Также посчитано среднее значение  $R^2$  по округам для каждой модели, для классической логистической модели составляет 0.78, для обобщенной логистической кривой - 0.97 и 0.95 - для модели Гомпертца. Сравнивая остальные коэффициенты моделей, отметим, что аппроксимации максимальных значений переболевших людей в классической модели и модели Гомпертца довольно близки по значению между собой. Вместе с тем коэффициент роста больше в модели Гомпертца, чем в

простой логистической модели примерно в 7 раз за счёт асимметрии решений в рамках модели Гомпертца.

Таким образом, для всех рассмотренных округов Российской Федерации обобщенная логистическая кривая является неплохой аппроксимацией натуральных данных, и эта модель описывает течение эпидемии в целом хорошо. Очевидно, что для всех рассмотренных округов и кривая Гомпертца неплохо описывает натурные данные ( $R^2 \geq 0.91$ ), но, однако, она уступает обобщенной логистической модели.



**Рис. 1. Связь количества заболевших в сутки ( $K$ ) и общего числа заболевших ( $N$ ) для округов РФ в рамках трех моделей: серый – простая логистическая модель, черный – обобщенная логистическая и светло-серым представлена модель Гомпертца**

Таблица 1

## Коэффициенты модели и коэффициент детерминации для трёх аппроксимаций

Модель	Простая логистическая			Обобщенная логистическая					Гомпертца		
	Округ	$N_{\infty}$	$r$	$R^2$	$N_{\infty}$	$r$	$\alpha$	$\beta$	$R^2$	$N_{\infty}$	$r$
Центральный	385000	0.07	<b>0.756</b>	697200	0.15	0.99	4.37	<b>0.945</b>	403200	0.56	<b>0.919</b>
Приволжский	160800	0.04	<b>0.765</b>	558900	2.77	0.63	5.58	<b>0.984</b>	181700	0.28	<b>0.953</b>
Уральский	173900	0.04	<b>0.879</b>	175200	4.1	0.56	0.67	<b>0.98</b>	201400	0.26	<b>0.97</b>
Южный	106000	0.05	<b>0.821</b>	199800	2.23	0.64	2.31	<b>0.975</b>	124700	0.3	<b>0.963</b>
Северо-Кавказский	106000	0.04	<b>0.715</b>	136400	5.9	0.53	1.23	<b>0.955</b>	119100	0.3	<b>0.925</b>
Сибирский	98640	0.04	<b>0.917</b>	108300	1.13	0.68	0.94	<b>0.976</b>	117600	0.26	<b>0.974</b>
Дальневосточный	118100	0.04	<b>0.722</b>	500000	2.23	0.63	6.52	<b>0.973</b>	136400	0.25	<b>0.939</b>
Северо-Западный	158800	0.04	<b>0.699</b>	292600	4.09	0.58	2.38	<b>0.969</b>	177900	0.26	<b>0.933</b>

Кроме Федеральных округов Российской Федерации анализ в рамках трех простейших моделей проведен также для всех 85 регионов РФ. В ходе исследования, выявлено пять регионов для которых коэффициент детерминации обобщенной логистической модели и модели Гомпертца различается не более чем на одну сотую (табл. 2).

Таблица 2.

## Сравнение коэффициента детерминации для двух аппроксимаций

Модель	Обобщенная логистическая	Гомпертца
Регион	$R^2$	$R^2$
Свердловская	<b>0.729</b>	<b>0.722</b>
Челябинская	<b>0.769</b>	<b>0.768</b>
Калмыкия	<b>0.502</b>	<b>0.503</b>
Иркутская	<b>0.845</b>	<b>0.843</b>
Калининградская	<b>0.665</b>	<b>0.66</b>

### Заключение

В данной работе для всех округов Российской Федерации произведен подбор наилучшей аппроксимации динамики распространения «первой волны» коронавирусной инфекции в рамках трех различных моделей (классическая и обобщенная логистические модели и модель Гомпертца). По результатам аппроксимации можно сделать вывод, что обобщенная логистическая модель является неплохой моделью для аппроксимации данных о заболевших среди представленных моделей. Стоит отметить также, что модель Гомпертца аппроксимирует данные с близкой точностью, хотя имеет меньшее количество подбираемых параметров.

Исследование выполнено при финансовой поддержке РФФИ и [CNPq, andNSFC] в рамках научного проекта № 20-51-80004», а также при поддержке гранта Президента РФ по государственной поддержке ведущих научных школ РФ НШ-70.2022.1.5.

### Библиографический список

1. **Pelinovsky, E.** Logistic equation and COVID-19 /Pelinovsky E. A., Kurkin A., Kurkina O., M., Kokoulina, M., Epifanova//Chaos, Solitons & Fractals. – 2020. – Т. 140. – С. 110241. DOI: 10.1016/j.chaos.2020.110241
2. **Pelinovsky, E.** Gompertz model in COVID-19 spreading simulation /Pelinovsky, E., Kirillin, M., Kokoulina, M., Epifanova, A., Kurkin A., Kurkina O., Tang M., Macau E.// Chaos, Solitons & Fractals. – 2022. – Vol. 154. – P. 111699. – DOI 10.1016/j.chaos.2021.111699.
3. **Кокоулина, М. В.** Анализ динамики распространения коронавируса с помощью обобщенной логистической модели / Кокоулина М. В., Епифанова А. С., Пелиновский Е. Н., Куркина О. Е., Куркин А. А. // Труды НГТУ им. Р.Е. Алексеева. – 2020. – № 3(130). – С. 28-41. – DOI 10.46960/1816-210X\_2020\_3\_28.
4. **Абрамов, С.М.** Об анализе, моделировании и прогнозе статистики эпидемии коронавируса/ Абрамов С. М., Травин С. О. //Цифровая экономика. – 2020. – №. 2. – С. 5-19. DOI: 10.34706/DE-2020-02-01
5. **Verhulst, P.** Notice sur la loi que la population suit dans son accroissement /P.Verhulst //Corresp. Math. Phys. – 1838. – Т. 10. – С. 113-126. DOI: 10.12691/jbms-1-2-1
6. **Carletti, T.** COVID-19: The unreasonable effectiveness of simple models / Carletti T., Fanelli D., Piazza F. //Chaos, Solitons & Fractals: X. – 2020. – Т. 5. – С. 100034. DOI:10.1101/2020.05.26.20110957
7. **Consolini, G.** A stretched logistic equation for pandemic spreading /Consolini G., Materassi M. //Chaos, Solitons & Fractals. – 2020. – Т. 140. – С. 110113. DOI:10.1016/j.chaos.2020.110113
8. **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
9. **Brilhante, M. F.** Modeling risk of extreme events in generalized Verhulst models / M. F. Brilhante, M. I. Gomes, D. Pestana // REVSTAT–Statistical Journal. – 2019. – V. 17. – № 2. – P. 145-162
10. **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.

**СВЕРХШИРОКОПОЛОСНЫЙ ИСТОЧНИК ОДНОРОДНОГО ЭЛЕКТРОМАГНИТНОГО  
ПОЛЯ ДИАПАЗОНА СВЧ**

**Е.П. Тимофеев**

ORCID:e-mail: [e.p.timofeev@icloud.com](mailto:e.p.timofeev@icloud.com)

Нижегородский государственный технический университет им. Р.Е. Алексеева

Рассмотрены источники однородного электромагнитного поля (ИО ЭМП) с целью воспроизведения максимально однородного поля в рабочей зоне, в широкой полосе частот. Предложена и разработана конструкция ИО ЭМП, обеспечивающая избирательное подавление волн высших типов. Представлены технические решения, позволяющие значительно расширить частотный диапазон одноволнового режима работы ИО ЭМП и увеличить область однородного поля. Получены новые конструкции и ИО ЭМП, имеющие существенно превосходящие характеристики аналогичных полеобразующих структур и позволяющие решать задачи ЭМС с меньшей погрешностью.

**Ключевые слова:** источник однородного электромагнитного поля (ИО ЭМП), экранированная полосковая линия, волны высших типов, радиопоглощающий материал (РПМ), одноволновый режим работы.

**ULTRA-WIDEBAND SOURCE OF A HOMOGENEOUS ELECTROMAGNETIC FIELD  
OF THE MICROWAVE RANGE**

**E.P. Timofeev**

Nizhniy Novgorod state technical university n.a.R.E. Alekseev

The sources of a homogeneous electromagnetic field (IO EMF) are considered in order to reproduce the most homogeneous field in the working area, in a wide frequency band. The design of the IO EMF, which provides selective suppression of higher types of waves, is proposed and developed. Technical solutions are presented that significantly expand the frequency range of the single-wave mode of operation of the IO EMF and increase the area of the homogeneous field. New designs and EMF IO have been obtained that have significantly superior characteristics of similar field-forming structures and allow solving EMC problems with less error.

**Keywords:** a source of a homogeneous electromagnetic field (IO EMF), a shielded strip line, waves of higher types, radio-absorbing material (RPM), single-wave mode of operation.



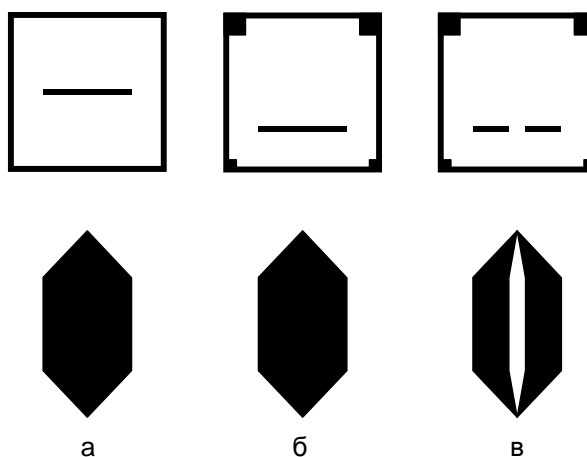
## Введение

Экранированные источники однородного электромагнитного поля (ИО ЭМП), широко используются при решении задач электромагнитной совместимости, сертификации радиоэлектронного оборудования, медицинских исследованиях [1].

В данной работе рассматриваются ИО ЭМП, выполненные на основе отрезка крупномасштабной экранированной полосковой линии (ЭПЛ) (переходящей с обеих сторон в согласующие конуса, подключенные соответственно к генератору на входе и согласованной нагрузке на выходе) с широким внутренним проводником, расположенным в экране либо симметрично (рис. 1а), либо асимметрично, (рис. 1 б, в) [1].

Недостатками данных источников являются:

- ограничение частотного диапазона одноволнового режима работы, вызванное возникновением волн высших типов, первыми из которых являются волны типа Н;
- относительно высокий уровень неоднородности ЭМП (определяемый по критериям [2]) в рабочей зоне, что естественно уменьшает ее размеры и снижает точность и достоверность поверки в ИО ЭМП.



а) – симметричный; б) – асимметричный; в) – проводник с продольной щелью

**Рис. 1. Варианты ИО ЭМП с различными внутренними проводниками**

В данной работе представлены технические решения (рис. 1 б, в), позволяющие уменьшить отмеченные выше недостатки и существенно улучшить характеристики ИО ЭМП.

Математические модели данных структур, рис. 1, (позволяющие рассчитывать погонные параметры, спектр критических частот волн высших типов, структуру ЭМП, области однородного ЭМП и другие характеристики), разрабатывались и анализировались в ряде работ авторов [1-3].

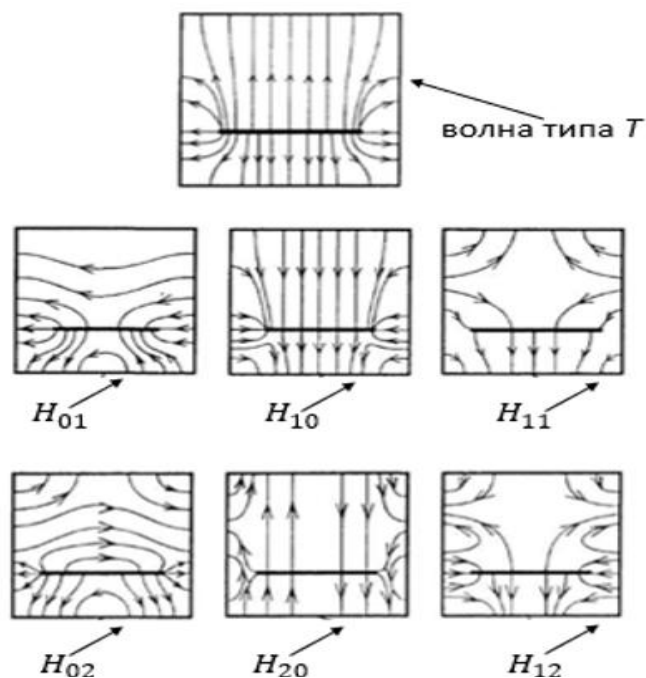
### Расширение рабочего диапазона частот ИО ЭМП (анализ, эксперимент)

В данной работе рассматривается структура, позволяющая значительно расширить диапазон одноволновой работы ИО ЭМП (рис. 1,б). Расширение этого диапазона возможно либо за счет правильного выбора геометрических размеров экранированной полосковой линии, позволяющего повысить критические частоты первых высших типов ( $H_{01}$ ;  $H_{10}$ ) волн, либо за счет их эффективного подавления.

Анализ спектра волн высших типов показал, что при смещении внутреннего проводника из максимума электрического поля волны  $H_{01}$  (симметричное расположение проводника) ее критическая частота повышается, расширяя тем самым частотный диапазон одноволнового режима работы.

Это явилось основанием для построения ИО ЭМП на основе несимметричной полосковой линии. Сказанное выше демонстрируется на рис.2, где для несимметричной экранированной линии квадратного сечения приведены результаты расчета структуры электрического поля основной ( $T$ -волны) и  $H$ -волн высших типов.

Дальнейшее расширение частотного диапазона ИО ЭМП возможно при эффективном подавлении волн высших типов без заметного влияния на основную волну.

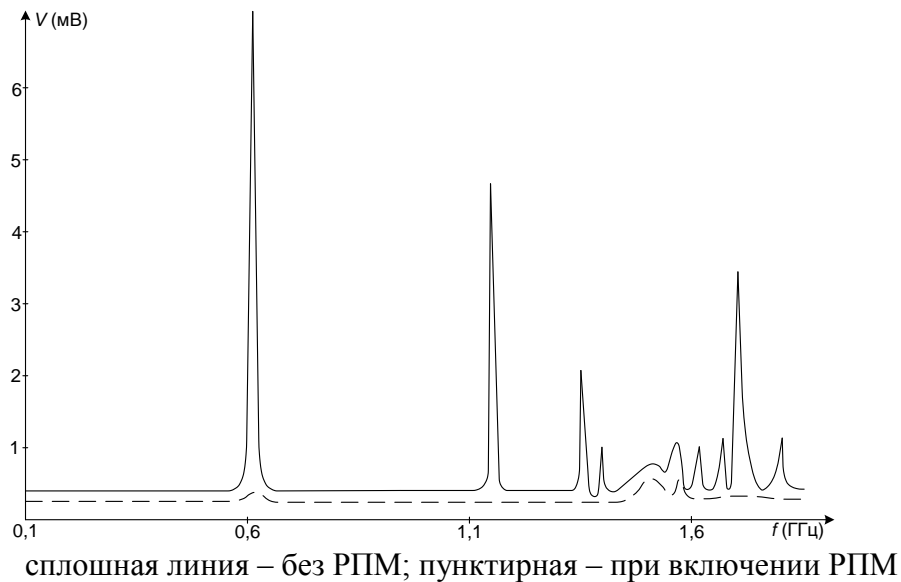


**Рис.2. Структуры электрического поля основной T-волны и H-волн высших типов**

Такая задача избирательного поглощения только волн высших типов решена на основе анализа структуры основной и высших типов волн и определения места размещения поглощающего материала, его конфигурации и размеров. Структура поля основной T-волны такова, что электрическое поле имеет максимальное значение вблизи щели и практически отсутствует в углах экрана. Структура поля волн высших типов имеет более сложный вид. Для них характерны вариации электрического поля на металлических поверхностях, образование петель электрического поля при изменении параметров экранированной полосковой линии. Для ряда волн типа-H такие петли образуются и в углах экрана, увеличивая тем самым концентрацию электрического поля в этих местах.

Помещенный в эти углы объемный радиопоглощающий материал (РПМ) квадратного сечения, эффективно взаимодействуя с электрическим полем волн высших типов, поглощает их (подавляет резонансы), практически не влияя на распространение основной T-волны. В результате диапазон одноволновой работы ИО ЭМП значительно расширяется, увеличиваясь как минимум в 3 раза. Экспериментально определено, что при поглотителе с оптимальными размерами [1] затухание основной T-волны увеличивается всего лишь на (5...7)% в то время как для H-волн высших типов – в (10...20) раз, обеспечивая практически их полное подавление. Этот эффект иллюстрируется на рис. 3, где показана зависимость выходного напряжения калибруемой антенны от частоты в реальном ИО ЭМП на основе ЭПЛ.

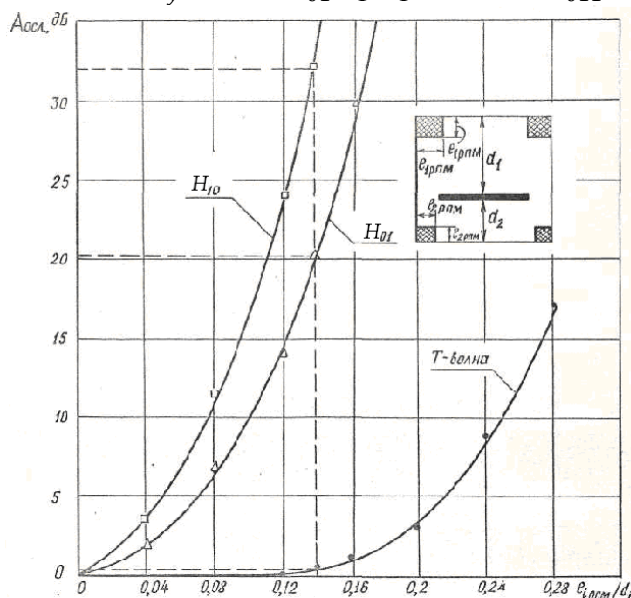
При выборе оптимального радиопоглощающего материала (РПМ), использовались экспериментальные исследования. В качестве объемного РПМ применялась смесь мелкодисперсного карбонильного железа с твердеющими заполнителями (полистирол, эпоксидная смола). Из анализа картин силовых линий поля  $E$  волн  $H_{01}, H_{10}, H_{11}$  и физических соображений следует, что наиболее оптимальной формой поперечного сечения РПМ, размещаемого во внутренних углах экрана является квадратная форма. Размеры поперечного сечения РПМ определялись на основе компромиссного решения между достижением достаточного уровня подавления резонансов  $H_{mn}$  - волн и предельной величиной ослабления T-волны в заданном диапазоне частот.



**Рис. 3. Резонансы волн высших типов**

На рис. 4 представлены графики зависимости ослабления компонент  $E_y$  ЭМП  $T$ -волны,  $H_{01}$ - волны и  $H_{10}$  - волны от размеров поперечного сечения РПМ –  $l_{ipnm}/d_i$ . Измерения ослабления  $T$ -волны проводились на частоте 200 МГц, волны  $H_{01}$  - на резонансной частоте  $f_{p011} = 222.5$  МГц и волны  $H_{10}$ ,- на резонансной частоте  $f_{p101} = 347.6$  МГц.

В соответствии с рис.4, за допустимую величину ослабления  $E_y$  - компонента  $T$ -волны принята величина равная 5% (0,4дБ). При этом размеры поперечного сечения РПМ не должны превышать  $l_{ipnm} \leq 0,14d_i$ . При этом, ослабление компоненты  $E_y$  волны  $H_{01}$  при резонансе  $H_{011}$  составляет порядка 20 дБ и компоненты  $E_y$  волны  $H_{10}$  при резонансе  $H_{011}$  - порядка 32 дБ.



**Рис.4. Зависимость ослабления  $A_{осл}$  компонент  $E_y$  волн  $T$ ,  $H_{10}$ ,  $H_{01}$  от размеров поперечного сечения РПМ**

Верхняя частота рабочего диапазона ИО ЭМП с РПМ может быть увеличена приблизительно в три раза, по сравнению с верхней частотой ИО ЭМП без РПМ.

Результаты измерений частотной зависимости  $T$ -волны,  $H_{01}$  -волны и  $H_{10}$  - волны в ИО ЭМП с РПМ позволяют сделать вывод, что РПМ, размещенный во внутренних углах экрана с  $l_{ipnm} \leq 0,14d_i$  размерамине оказывает заметного влияния на  $T$ -волну.

## Увеличение рабочей области однородного поля широкополосного ИО ЭМП

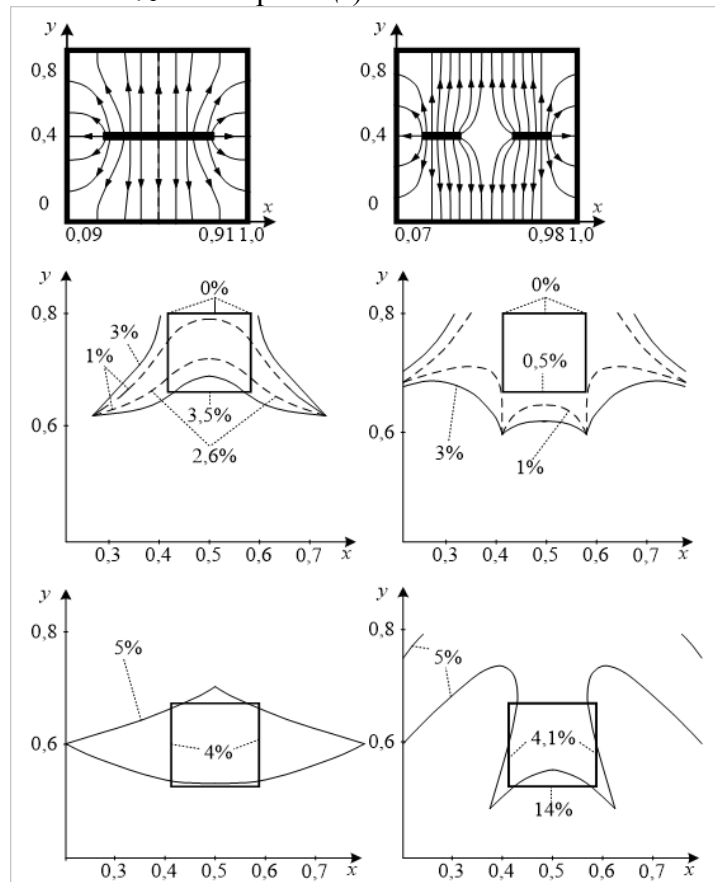
При проектировании ИО ЭМП важной задачей является воспроизведение максимально однородного поля в рабочей зоне. Такая задача увеличения однородности ЭМП в рабочей зоне решена за счет использования в качестве внутреннего полоска проводника с широкой продольной щелью по всей длине (рис.1,в). Важную роль при этом имеет правильный выбор параметров внутреннего проводника (размеров полосков и ширины продольной щели). Задача оптимального выбора его параметров решена на основе анализа распределения ЭМП по поперечному сечению многопроводной экранированной полосковой линии, выполненного с помощью разработанного алгоритма [4]. При этом границы области однородного ЭМП определялись с помощью следующего критерия:

$$\delta\% = \left| 1 - [E_{yik}(x = x_i, y = y_k) / E_{y0}(x = 0, y = d/2)] \right| \cdot 100\%,$$

который указывает границы  $(x_i, y_k)$  области однородного ЭМП, где компонента  $E(x_i, y_k)$  изменяется относительно значения электрического поля в центре рабочей зоны ( $E_{y0}(x = 0, y = d/2, Ex0x=0, y=0)$ ) не более чем на заданную величину  $\delta\%$ .

Для ИО ЭМП, со сплошным широким внутренним проводником (рис.1а,б) это отклонение составляет не менее  $\delta\% = 5\%$ . Это связано с тем, что значение компоненты  $E_y$  всегда уменьшается по мере прохождения от внутреннего проводника (по оси ординат) при любом размещении рабочей зоны.

При использовании внутреннего проводника с широкой щелью (рис.1, в) однородное ЭМП формируется за счет боковых проводников. При этом компонента  $E_y$  возрастает (по оси ординат) (в отличие от рис.1, а) достигая максимального значения незначительно спадая у поверхности экрана. Подобное поведение компоненты  $E_y$  позволяет сформировать практически однородное ЭМП. Неоднородность ЭМП  $\delta\%$  на границе рабочей зоны не превышает 0,5%, что в (8...10) раз меньше, чем в случае структуры на рис. 1,а. Кроме того при данном техническом решении увеличивается область однородного ЭМП, что позволяет расширить рабочую зону не менее, чем в 3 раза (при заданном значении  $\delta\%$  на ее границе).



**Рис. 5. Сравнительный анализ ИО ЭМП**

На рис. 5 представлены результаты сравнительного анализа различных вариантов полеобразующих структур, поясняющие процесс формирования областей однородного ЭМП. При этом размеры внешнего экрана для всех вариантов оставались неизменными. Волновое сопротивление  $W = 50 \text{ Ом}$ . Здесь же представлены результаты расчета (по введенному выше критерию) областей однородного ЭМП для анализируемых ЭПЛ при размещении рабочей зоны (квадратного сечения) у поверхности экрана (а) и в центре верхней полости ЭПЛ (б). На границе рабочей зоны (в ее характерных точках) и областей однородного ЭМП приведены значения неоднородности поля  $\delta\%$ .

### Выводы

1. Предложена и разработана новая конструкция ИО ЭМП, обеспечивающая избирательное подавление волн высших типов.
2. Показано, что введение радиопоглощающего материала расширяет рабочий частотный диапазон одноволнового режима работы ИО ЭМП в 3 раза.
3. Показано, что при использовании широкого внутреннего проводника с продольной щелью достигнуто увеличение области однородного ЭМП и расширение рабочей зоны ИО ЭМП в 3 раза, позволяющее увеличить размеры поверяемых антенн.
4. Получены новые конструкции ИО ЭМП, имеющие существенно превосходящие характеристики аналогичных полеобразующих структур, позволяющие решать задачи ЭМС с меньшей погрешностью.

### Библиографический список

1. Тимофеев, Е.П. Патент РФ 1663586 МПК G01R33/02 Устройство для калибровки измерителей напряженности электромагнитного поля // А.И.Ермолаев, В.А.Пронин, Е.П.Тимофеев. Оpubл. 15.07.91, Бюл. №26.
2. Тимофеев, Е.П. Выбор границ области однородного поля при аттестации электрически малых антенн ТЕМ-камерах / С.Н.Аржанов, В.А.Пронин, Е.П.Тимофеев, В.П.Хилов // Техника средств связи. Сер.РТ, 1982. Вып.7(46). С.46-53.
3. Тимофеев, Е.П. Исследование структуры полей собственных волн источников однородного поля /А.И.Ермолаев, В.А.Пронин, Е.П.Тимофеев // Проблемы радиоизмерительной техники: Тез.докл. Всесоюз. симп., ГНИПИ Горький, 1989. С.116-117.
4. Илларионов, Ю.А. Расчёт элементов полосковых трактов СВЧ /Ю.А. Илларионов, В.Я.Сморгонский, Е.П.Тимофеев // Расчёт элементов полосковых трактов СВЧ: учебное пособие. Горький, Горьковск.политех.ин-т, 1985. 102с.

**АЛГОРИТМЫ АНАЛИЗА И ОБРАБОТКИ ГЕОИНФОРМАЦИИ ДЛЯ ПОДДЕРЖКИ  
ПРИНЯТИЯ РЕШЕНИЙ В ЗАДАЧАХ СИНТЕЗА СТРУКТУРЫ СИСТЕМ  
СПУТНИКОВОЙ РАДИОНАВИГАЦИИ**

**Белов Ю.Г.<sup>1</sup>**

ORCID: 0000-0002-2753-119X e-mail: bel266@nntu.ru

**Бирюков В.В.<sup>1</sup>**

ORCID: 0000-0002-6003-0450 e-mail: birukovvv@mail.ru

**Бисярин М.Н.<sup>2</sup>**

ORCID: 0000-0003-1052-7247 e-mail: maksimbis@mail.ru

**Котылёва Е.М.<sup>1</sup>**

ORCID: 0000-0002-5198-7795 e-mail: e.kotyleva@bk.ru

**Малахов В.А.<sup>1</sup>**

ORCID: 0000-0001-5902-6908 e-mail: mr.vasmal@mail.ru

**Раевская Ю.В.<sup>1</sup>**

**Раевский А.С.<sup>1</sup>**

ORCID: 0000-0001-8678-0949 e-mail: raevsky\_as@mail.ru

**Седаков А.Ю.<sup>2</sup>**

**Титаренко А.А.<sup>1</sup>**

**Щербаков В.В.<sup>1</sup>**

ORCID: 0000-0002-7744-0763 e-mail: sherbakovvv@gmail.com

<sup>1</sup>Нижегородский государственный технический университет имени Р.Е. Алексева

<sup>2</sup>Филиал РФЯЦ-ВНИИЭФ «НИИИС им. Ю.Е. Седакова»

Описывается разработка математической модели для реализации пространственной, селекции радиочастотных помех многоэлементной антенной решеткой (АР) аппаратуры спутниковой навигации (АСН) путём формирования максимумов диаграммы направленности (ДН) в направлении на навигационные космические аппараты (НКА), и алгоритма, реализующего данную математическую модель. Разработан алгоритм пространственной селекции помех, обеспечивающий для заданного момента времени: определение доступной группировки НКА ГЛОНАСС и GPS в зависимости от текущего положения НКА и угловой ориентации АР; расчет и построение (визуализацию) ДН многоэлементной АР для каждого видимого НКА; расчет и построение (визуализацию) результирующей ДН многоэлементной АР. Полученные результаты использованы для создания системы формирования ДН АР, необходимую пространственную селекцию радиочастотных помех.

**Ключевые слова:** Алгоритм пространственной селекции помех, синтез диаграммы направленности, антенная решетка, спутниковая навигация.

**ALGORITHM AND SOFTWARE FOR ANALYSIS AND PROCESSING OF INFORMATION  
RETRIEVED FROM THE INTERNET FOR DECISION SUPPORT EQUIPMENT FOR  
SATELLITE NAVIGATION HARDWARE STRUCTURE SYNTHESIS**

**Y.G. Belov<sup>1</sup>, V.V. Biryukov<sup>1</sup>, M.N. Bisyarin<sup>2</sup>, E.M. Kotylyova<sup>1</sup>, V.A. Malakhov<sup>1</sup>, Y.V. Raevskaya<sup>1</sup>, A.S. Raevsky<sup>1</sup>, A.Y. Sedakov<sup>2</sup>, A.A. Titarenko<sup>1</sup>, V.V. Shcherbakov<sup>1</sup>**

<sup>1</sup>Nizhny Novgorod State Technical University n.a. R.E. Alekseev

<sup>2</sup>Branch of FSUE RFNC-VNIIEF «NIIS n.a. U.E. Sedakov»

Development of mathematical model for realization of spatial, radio-frequency interference selection by multielement antenna array of satellite navigation equipment by formation of maximums of directivity diagram in the direction to navigation spacecrafts, and algorithm realizing this mathematical model are described. The algorithm of spatial interference selection is developed in the high-level programming language C# in the integrated development environment MS Visual

Studio. The algorithm provides for a given point in time: to determine the available grouping of navigation spacecraft (NS) GLONASS and GPS, depending on the current position of the NS and the angular orientation of the antenna array (AA), taking into account the trajectory; calculation and construction (visualization) of multi-element AA for each visible NSA; calculation and construction (visualization) of the resulting multi-element AA for each visible navigation spacecraft. The use of an advanced antenna system also makes it possible to implement an algorithm for calculating the autocompensator of interference, which allows the formation of minima of the directional pattern of the adaptive phased array in the directions towards interference sources, while maintaining the same level in the other directions. This method does not require a priori information about the angular location of the useful signal source and interference.

The results obtained in the process of work can be used to create a system of formation of the antenna array radiation pattern, which provides the necessary spatial selection of radio-frequency interference. The implementation of the algorithm consists in converting the ephemeris data of GLONASS and GPS satellites to geographical coordinates (longitude and latitude), followed by their recalculation into the coordinate system of the antenna array, which in the general case moves along a given trajectory. By changing the phase distribution of the antenna array (whose position is known), the directivity pattern is achieved so that the maximum of the main lobe is directed towards the visible satellite, a necessary spatial selection of radio-frequency interference. A description of the interface of the developed program and peculiarities of work with it are given.

**Keywords:** Spatial interference selection algorithm, radiation pattern synthesis, antenna array, satellite navigation.

## Введение

В настоящее время большое распространение получили спутниковые радионавигационные технологии. На их основе создаются навигаторы различного назначения, от геодезической и специальной аппаратуры до бытовых навигаторов, а также аппаратура частотно-временной синхронизации [1].

В то же время ключевым недостатком спутниковых радионавигационных систем является низкая помехоустойчивость, обусловленная тем, что высота орбиты навигационных спутников составляет порядка 20000 км. Большое расстояние от навигационных спутников GPS, ГЛОНАСС, Galileo и других до аппаратуры потребителей приводит к тому, что спектральная плотность мощности полезных сигналов становится меньше, чем мощность шума в антенне. Малая мощность принимаемых спутниковых навигационных сигналов приводит к тому, что помехи естественного или искусственного происхождения, которые соизмеримы по мощности с полезным сигналом, уже могут представлять опасность для целостности навигационных сообщений [2].

В случае приема навигационных сигналов на фоне мощной электромагнитной помехи значительное увеличение отношения «сигнал/помеха» может быть достигнуто только путем пространственной фильтрации сигналов при помощи автокомпенсаторов помех. В основе автокомпенсаторов лежит многоантенная система – антенная решетка (АР) с диаграммообразующей схемой (ДОС), с помощью которых формируется диаграмма направленности (ДН) АР. Существуют два принципиальных способа управления ДН АР. Первый состоит в том, чтобы поддерживать уровень принимаемой мощности со всех направлений на одинаковом уровне, но минимизировать её в направлении нежелательного сигнала. Второй способ выражается в максимизации принимаемой мощности со стороны источника требуемого сигнала и минимизации принимаемой мощности с остальных направлений.

Первый способ не требует знать положение источников полезных сигналов. Управление ДН АР может производиться только на основе сведений о помеховой обстановке. Для АР диапазона спутниковых навигационных систем критерием управления ДН является подавление всех сигналов в рабочем диапазоне частот, чья мощность превышает мощность шума в антенных элементах.

Для случая приема навигационных сигналов второй способ управления ДН АР требует дополнительно информацию о положении навигационных космических аппаратов (НКА),

поскольку для решения навигационной задачи необходим прием сигналов минимум от четырех спутников. Также минимизация мощности сигналов с направлений, не соответствующих полезным сигналам, может быть недостаточной для устранения помехи из входного сигнала навигационного приемника. Информация о положении НКА глобальных навигационных спутниковых систем (ГНСС) может быть получена путем преобразования сведений из альманахов и эфемерисов (набора сведений о текущем состоянии навигационной системы) в координаты спутников в сферической системе координат, связанной с фазовым центром АР. Второй недостаток может быть устранен подбором специального неравноамплитудного питания элементов АР, такого, чтобы уровень максимального бокового лепестка ДН не превышал уровень собственных шумов навигационного приемника.

Широкое применение для решения задачи увеличения помехоустойчивости радиоэлектронной аппаратуры получили адаптивные антенные решетки. Под антенной решеткой понимается набор антенных элементов, расположенных в пространстве таким образом, что суммарная ДН АР отличается от ДН каждого элемента по отдельности. Процесс формирования ДН АР заключается в домножении принятого сигнала каждого антенного элемента на комплексный весовой коэффициент, который изменяет амплитуду и фазу сигнала, с последующим суммированием сигналов со всех антенных элементов. Адаптивной антенной решеткой (ААР) такая система будет в том случае, если набор весовых коэффициентов изменяется с течением времени, адаптируясь тем самым к изменению сигнально-помеховой обстановки. Управление амплитудой и фазой комплексных весовых коэффициентов позволяет сформировать направление и ширину главного лепестка (одного или нескольких), а также минимумы в ДН ААР [3].

Особенность обработки навигационных радиосигналов с использованием ААР заключается в том, что полезный сигнал по мощности расположен на несколько децибел ниже уровня шумов. Поэтому любой сигнал выше уровня шумов можно рассматривать как помеху. Именно на этом свойстве основан первый способ управления ДН ААР, о котором говорилось ранее. Однако данный подход имеет ряд существенных недостатков. Для эффективной работы необходимо, чтобы мощность полезного сигнала на входе антенных элементов была много меньше уровня шума или отсутствовала вообще. Другим недостатком является невозможность выставить максимум ДН ААР в направлении радионавигационного сигнала. Также, в случае отсутствия сигнала помехи, весовые коэффициенты некоторых элементов АР будут равны нулю, тем самым уменьшая общий коэффициент усиления (КУ) АР [4].

Поэтому второй способ управления ДН ААР является более эффективным для решения навигационной задачи. Он также позволит увеличить уровень принимаемых навигационной аппаратурой сигналов.

Особенно важно при выборе данного способа формирования ДН ААР отслеживать характеристики каждого из лучей многолепестковой ДН, максимумы которой направлены на НКА, поскольку при отклонении луча ААР от нормали к решетке сильно меняется как форма ДН, так и другие радиотехнические параметры. Существует несколько способов наглядного представления (визуализации) ДН: пространственное изображение ДН и двумерное. В случае визуализации многолепестковой ДН двумерное представление не является наглядным, поскольку для информации о форме ДН необходимо минимум два сечения во взаимно перпендикулярных плоскостях для каждого луча многолепестковой ДН. Поэтому пространственное изображение ДН ААР для мониторинга основных её характеристик будет более информативным.

### **1 Алгоритм расчёта направления на НКА**

Исходными данными являются значения эфемерид из альманахов, приведённых в [5]. Сначала рассчитывается положение НКА в географической системе координат (широта и долгота). Поскольку эфемериды ГЛОНАСС и GPS различны [6,7], этот этап расчёта для них различается.



### 1.1 Расчёт долготы и широты НКА ГЛОНАСС

По известным значениям географической долготы восходящего узла и угла наклона орбиты находим угловую длину дуги, пройденной НКА к данному моменту времени. Из сферического треугольника  $ABC$  (рис. 1), образованного этой дугой ( $AB$ ) и дугами «экватора» ( $AC$ ) и «меридиана» ( $BC$ ), по теореме синусов находим широту НКА в данный момент времени в радианах  $\beta_r$  и в градусах  $\beta$ .

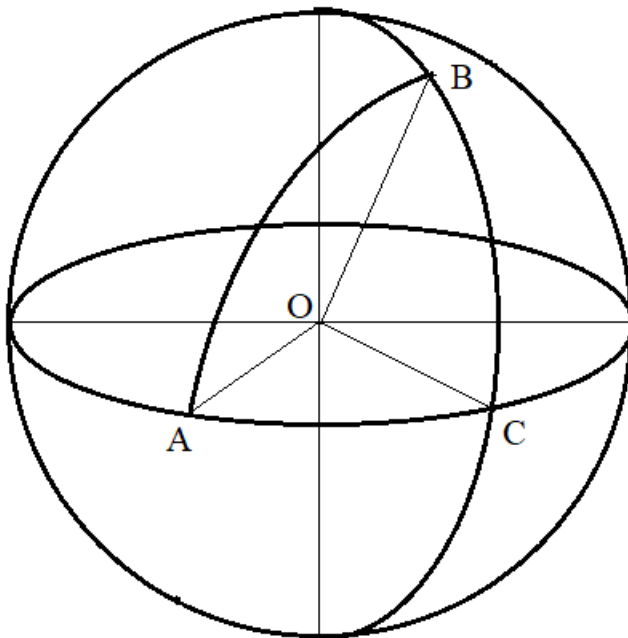


Рис. 1. Сферический треугольник  $ABC$ ,  $A$ – восходящий узел

$$\beta_r = \arcsin \left( \sin \left( i \frac{\pi}{180} \right) \sin \left( \frac{2\pi(t-T\Omega)}{T_{об}} \right) \right),$$

$$\beta = \beta_r \frac{180}{\pi}.$$

По теореме косинусов находим угловую длину экваториальной дуги. Добавляя её к значению географической долготы восходящего узла и учитывая поворот Земли за рассматриваемый промежуток времени, получаем географическую долготу НКА –  $\alpha$ .

$$\alpha = \text{sign}(\beta) \left[ \frac{\cos \left( \frac{2\pi(t-T\Omega)}{T_{об}} \right)}{\cos(\beta_r)} \right] \frac{180}{\pi} + L\Omega - (t-T\Omega)\omega_0 \frac{180}{\pi},$$

где  $\omega_0 = 7,292115 \cdot 10^{-5} \text{ c}^{-1}$  – угловая скорость вращения Земли.

Учитывая, что за рассматриваемый промежуток времени НКА может совершить несколько оборотов вокруг Земли, окончательно получаем

$$\alpha_0 = \alpha - \left( \text{tranc} \left( \frac{\alpha}{180} \right) - \text{tranc} \left( \frac{\alpha}{360} \right) \right) 360.$$

Полагая  $\alpha = \varphi$ ,  $\beta = 90^\circ - \Theta$ ,  $r = R_3 + h_{\text{НКА}}$ , получаем координаты НКА в сферической системе координат, в которой  $\Theta$  и  $\varphi$  находим из приведённых выражений,  $h_{\text{НКА}}$  – высота орбиты НКА,  $R_3$  – радиус земли для данной широты. Для разных широт радиус Земли различен [8], в программе он рассчитывается по формуле

$$R = \sqrt{\frac{(a^2 \cos(\beta))^2 + (b^2 \sin(\beta))^2}{(a \cdot \cos(\beta))^2 + (b \cdot \sin(\beta))^2}},$$

где  $a$  и  $b$  – большая и малая полуоси общеземного эллипсоида.

### 1.2 Расчёт долготы и широты НКА GPS

Используя данные эфемерид, находим угловую длину дуги орбиты, пройденной НКА к данному моменту времени:

$$\psi = m + \omega + \omega_s (t_s - t) \frac{180}{\pi},$$

где  $m$  – средняя аномалия,  $\omega$  – аргумент перигея,  $\omega_s$  – угловая скорость движения НКА по орбите.

$$\omega_s = \sqrt{\frac{\mu}{A^3 \cdot 10^9}},$$

где  $A$  – большая полуось орбиты НКА,  $\mu$  – геоцентрическая гравитационная постоянная.

Из сферического треугольника, образованного этой дугой и дугами «экватора» и «меридиана», по теореме синусов находим широту НКА в данный момент времени:

$$\beta_r = \arcsin \left[ \sin \left( i \frac{\pi}{180} \right) \cdot \sin \left( \psi \frac{\pi}{180} \right) \right],$$

$$\beta = \beta_r \cdot \frac{180}{\pi}.$$

По теореме косинусов находим угловую длину экваториальной дуги. Добавляя её к значению географической долготы восходящего узла и учитывая поворот Земли за рассматриваемый промежуток времени, получаем географическую долготу НКА:

$$\alpha = \text{sign}(\beta) \cdot \arccos \left( \frac{\cos \left( \frac{\psi \pi}{180} \right)}{\cos(\beta_r)} \right) \cdot \frac{180}{\pi} + L\Omega - (t_s + t) \omega_0 \cdot \frac{180}{\pi},$$

Далее учитывается, что за рассматриваемый промежуток времени НКА может совершить несколько оборотов вокруг Земли.

Полагая  $\alpha = \varphi$ ,  $\beta = 90^\circ - \Theta$ ,  $r = R_3 + h_{\text{НКА}}$ , получаем координаты НКА в сферической системе координат.

### 1.3 Переход к связанной системе координат

Переход к системе координат, связанной с аппаратурой спутниковой навигации (АСН), осуществляется в три этапа.

Сначала производим переход из сферической системы координат в декартову. Декартовы координаты АСН ( $x_A, y_A, z_A$ ) берем из данных траектории для выбранного момента времени. По ним находим широту  $La$  и долготу  $Lo$  АСН. Затем производим параллельный перенос системы координат в точку нахождения АСН. Вращая получившуюся систему координат на угол  $\gamma$  вокруг оси  $OZ$ , направляем ось  $OX$  на север. Затем поворачиваем вокруг  $OY$  на угол  $\delta$ , направляя ось  $OZ$  вверх по местной вертикали.

Учитываем рысканье, тангаж и крен. Для этого производим повороты системы координат на соответствующие углы (они вводятся из данных траектории). Переходя в сферическую систему координат, находим углы  $\Theta$  и ориентации антенны АСН на выбранный НКА.

В зависимости от ориентации АСН относительно горизонтальной поверхности часть углов обзора антенны может быть закрыта земной поверхностью (рис. 2). Для нахождения этих

углов рассмотрим кватернионы [8], описывающие повороты АСН на угол тангажа  $\alpha_T$  и угол крена  $\alpha_K$ :

$$Q_1 = \left\{ \cos \frac{\alpha_T}{2}; 0, \sin \frac{\alpha_T}{2}, 0 \right\};$$

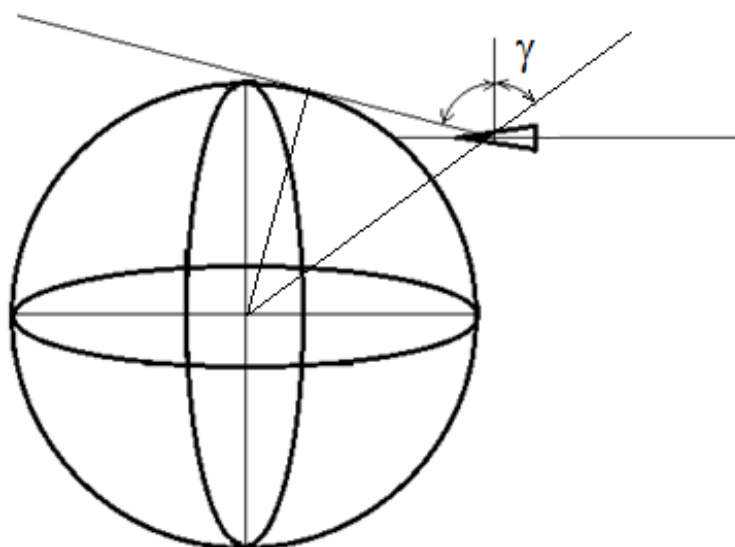
$$Q_2 = \left\{ \cos \frac{\alpha_K}{2}; \sin \frac{\alpha_K}{2}, 0, 0 \right\}.$$

Кватернион, соответствующий обоим поворотам, равен их произведению:

$$Q = Q_1 Q_2 = \left\{ \cos \frac{\alpha_T}{2} \cos \frac{\alpha_K}{2}; \cos \frac{\alpha_T}{2} \sin \frac{\alpha_K}{2}, \sin \frac{\alpha_T}{2} \cos \frac{\alpha_K}{2}, -\sin \frac{\alpha_T}{2} \sin \frac{\alpha_K}{2} \right\}.$$

Таким образом, нормальная ось системы координат, связанной с подвижной АСН, отклонена от направления местной нормали на угол

$$\gamma = 2 \arccos \left( \cos \frac{\alpha_T}{2} \cos \frac{\alpha_K}{2} \right).$$



**Рис. 2. Углы обзора антенны в зависимости от ориентации АСН относительно горизонтальной поверхности**

Этот угол определяет максимальное ограничение обзора поверхностью Земли. Для определения направления, в котором действует это ограничение, найдём направление проекции нормальной оси на горизонтальную поверхность после поворотов на углы тангажа и крена:

$$Q_Z = \{0; 0, 0, 1\};$$

$$Q_{Z'} = Q Q_Z Q^{-1} = Q Q_Z \tilde{Q} = \{0; \sin \alpha_T \cos \alpha_K, -\sin \alpha_K, \cos \alpha_T \cos \alpha_K\}.$$

Т.к. нормы данных кватернионов равны 1, обратный кватернион  $Q^{-1}$  равен сопряженному  $\tilde{Q}$ .

Таким образом, направление, в котором влияние земной поверхности максимально, определяется выражением

$$\varphi_m = \arctg 2(\sin \alpha_T \cos \alpha_K, -\sin \alpha_K).$$

При отклонении от  $\varphi_m$  на  $\pm \frac{\pi}{2}$  угол, ограничивающий обзор, уменьшается до нуля.

#### 1.4 Блок-схема программы расчета направления на НКА

Блок-схема программы, соответствующая алгоритму расчета направления на НКА представлена на рис. 3.

Программа пересчитывает эфемериды в долготу и широту НКА, затем переводит их в декартову систему координат, центр которой совпадает с центром земли, после систему параллельно переносят в место нахождения антенны. Полученную систему последовательно поворачивают вокруг  $OZ$  и  $OY$ , программа учитывает угловое движение: тангаж, рыскание и крен. Далее находятся координаты НКА в сферической системе и определяется его доступность для АСН.

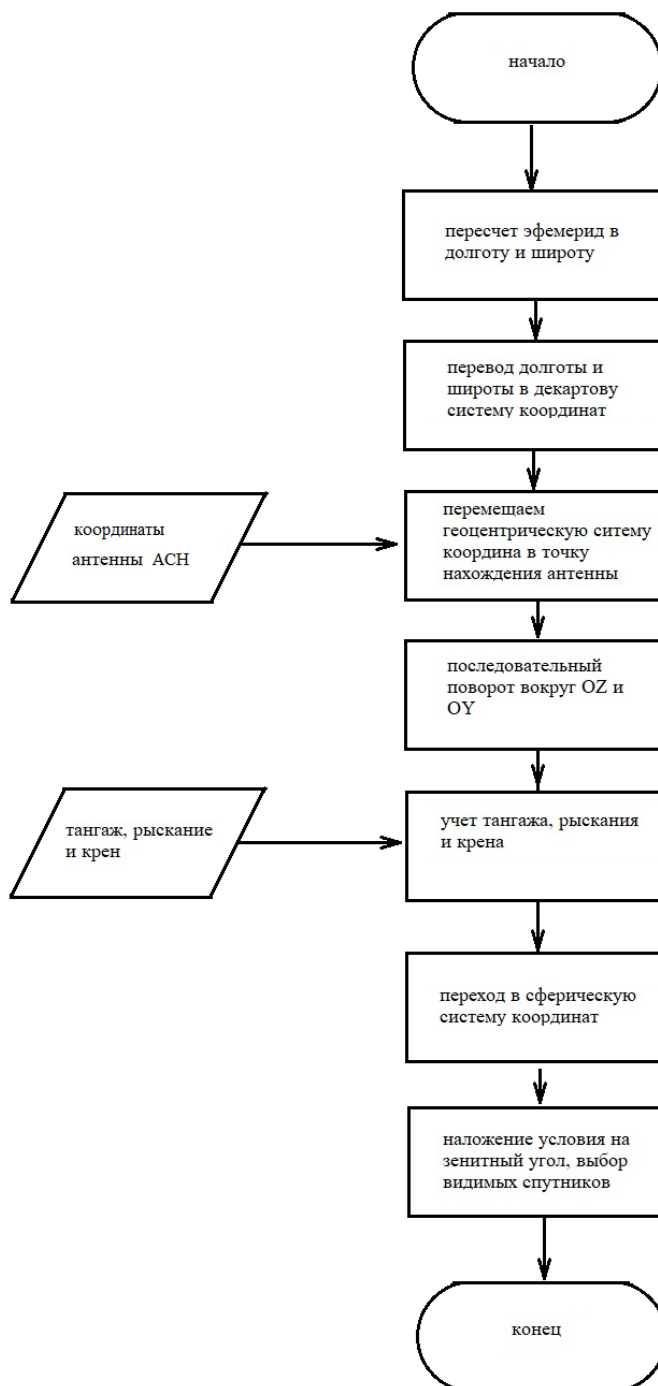


Рис. 3. Блок-схема программы, соответствующая алгоритму расчета направления на НКА

## 2 Расчет фазового распределения плоской квадратной антенной решетки

### 2.1 Получение расчетных соотношений

Задача: необходимо направить луч антенной решетки (АР) на навигационный космический аппарат (НКА), координаты которого в сферической системе, определяются через полярный угол  $\theta$  и азимутальный угол  $\varphi$  (рис. 4).

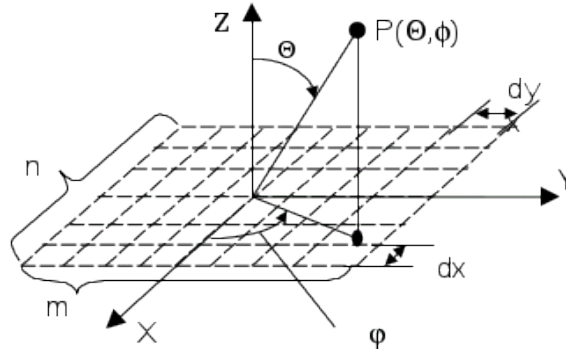


Рис. 4. Плоская антенная решетка в сферической системе координат

Рассмотрим наиболее простой случай – равноамплитудное питание:  $|I_n| \equiv 1$ .

Для того чтобы качать луч плоской АР в произвольном направлении, необходимо создать линейно нарастающее вдоль решетки фазовое распределение. В случае плоской АР это соответствует линейному фазовому распределению вдоль осей  $x$  и  $y$  соответственно. В таком случае амплитудно-фазовое распределение (АФР) будет иметь вид[9,10]:

$$I_n(x) = I_0 \exp[-j(n(x)-1)\Delta\Phi(x)];$$

$$I_n(y) = I_0 \exp[-j(n(y)-1)\Delta\Phi(y)],$$

где  $I_0 = 1$ ;  $\Delta\Phi(x)$ ,  $\Delta\Phi(y)$  – разность фаз между двумя любыми соседними излучателями;  $n(x)$ ,  $n(y)$  задают номера излучателей вдоль осей  $x$  и  $y$  соответственно.

Выражение для нормированного множителя направленности в сферической системе координат можно записать в виде:

$$F(\theta, \varphi) = \left| \frac{\sin \left[ \frac{N_x k d_x}{2} (\sin \theta \cos \varphi - \xi_x) \right]}{N_x \sin \left[ \frac{N_x k d_x}{2} \frac{(\sin \theta \cos \varphi - \xi_x)}{N_x} \right]} \frac{\sin \left[ \frac{N_y k d_y}{2} (\sin \theta \sin \varphi - \xi_y) \right]}{N_y \sin \left[ \frac{N_y k d_y}{2} \frac{(\sin \theta \sin \varphi - \xi_y)}{N_y} \right]} \right|.$$

Направление главного максимума излучения определится из равенств:

$$\xi_x = \sin \theta_0 \cos \varphi_0 = \frac{\Delta\Phi_x}{k d_x},$$

$$\xi_y = \sin \theta_0 \sin \varphi_0 = \frac{\Delta\Phi_y}{k d_y}.$$

$k$  – волновое число свободного пространства.

Эти формулы позволяют найти необходимые коэффициенты замедления фазовой скорости для ориентации главного максимума излучения в заданном направлении  $\theta_0$ ,  $\varphi_0$ .

### 2.2 Описание программы расчета

Алгоритм визуализации ДН АР был реализован на языке программирования С# [11-12], блок-схема которого представлена на рис.5.

На указанном рисунке:

1. Блок ввода данных. Исходными для расчета ДН плоской АР являются: направление на спутник, определяется через полярный угол  $\theta_0$  и азимутальный угол  $\varphi_0$ , число элементов АР  $N_x, N_y$  вдоль осей  $x$  и  $y$  соответственно, верхняя  $f_v$  и нижняя  $f_n$  частота диапазона навигационной системы, число  $n$ , определяющее точность построения ДН АР.



Рис. 5. Блок схема алгоритма визуализации ДН АР

2. Блок расчета основных параметров.

Изначально определяется средняя частота диапазона навигационной системы по формуле. По рассчитанной частоте определяется средняя длина волны  $\lambda_{cp}$ , затем волновое число  $k$ . Далее определяется шаг решетки. По введенным углам  $\theta_0$  и  $\varphi_0$  определяются коэффициенты замедления фазовой скорости возбуждения  $\xi_x$  и  $\xi_y$ .

3. Третий блок определяет точность построения ДН АР. Чем больше число  $n$ , тем точнее будет визуализирована ДН. Однако также время, затрачиваемое на ее построение, будет больше.

4. В четвертом блоке диапазоны изменения полярного и азимутального углов дискретизируются.

5. В пятом блоке пространственные сферические углы  $\theta_i, \varphi_j$  подставляются в выражение для множителя направленности плоской АР. Также осуществляется переход в декартовую систему координат, согласно формулам.

6. В шестом блоке по сформированному массиву точек осуществляется визуализация ДН плоской квадратной АР с помощью специальных средств программной среды разработки.

### 2.3 Учет направленных свойств элемента антенной решетки

В качестве излучающего элемента антенной решетки используются планарные излучатели. Наибольшее распространение получили пластины прямоугольной, круглой или эллиптической геометрических форм. Длина пластины составляет приблизительно половину длины волны. Возбуждение таких излучателей осуществляется с помощью коаксиального или полоскового фидера. В качестве диэлектрического основания используются диэлектрики с относительной диэлектрической проницаемостью от 2 до 10 с низкими потерями –  $\text{tg}\delta \approx 10^{-3}$ . Толщина диэлектрической пластины составляет  $(0,1-0,01)\lambda$ . Диаграмма направленности таких излучателей осесимметричная с коэффициентом усиления 7-9 дБм и с шириной главного лепестка около 65 градусов. Такая диаграмма направленности хорошо аппроксимируется с помощью функции  $p \cos^q \theta$ . Амплитудный коэффициент  $p$  легко находится при задании коэффициента усиления излучающего элемента антенной решетки – КУ (дБ) из выражения  $\text{КУ (дБ)} = 20 \lg(p)$ . Степенной показатель  $n$  у тригонометрической функции  $\cos \theta$  определяется при задании ширины диаграммы направленности излучающего элемента антенной решетки  $2\Delta\theta$  из выражения  $\cos^q \Delta\theta = \frac{1}{\sqrt{2}}$ .

Полученное таким образом выражение для диаграммы направленности излучающего элемента антенной решетки подставляем в выражение теоремы перемножения и получаем результирующее выражение для диаграммы направленности антенной решетки с заданным излучающим элементом.

В разработанной программе предусмотрен интерфейс, представляющий из себя два окна для ввода данных: коэффициента усиления в дБ и ширины ДН элемента АР по уровню –3 дБ в градусах.

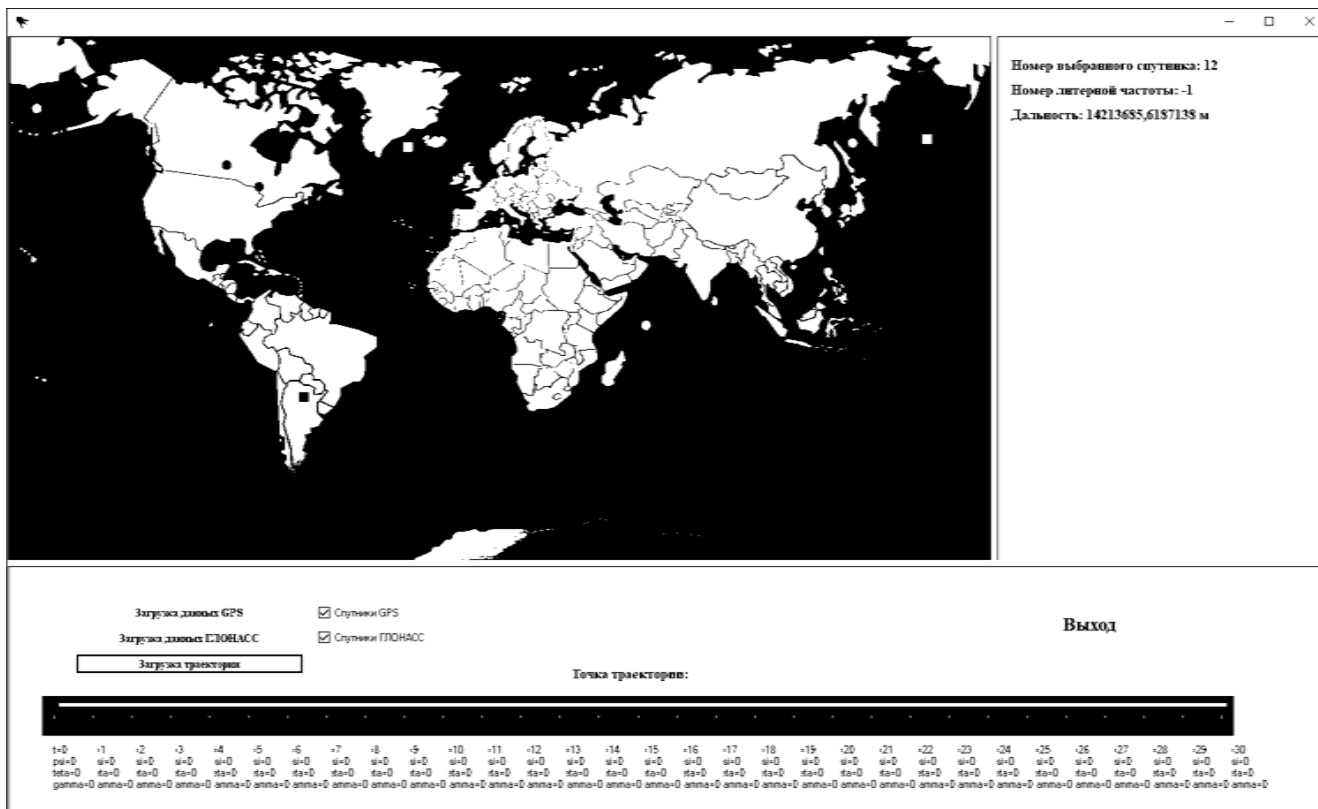
## 3 Интерфейс программы пространственной селекции помех для аппаратуры спутниковой навигации с многоэлементной антенной решеткой

Программа пространственной селекции помех для аппаратуры спутниковой навигации с многоэлементной антенной решеткой написана на языке программирования высокого уровня С# [7-11] в интегрированной среде разработки MS Visual Studio. Программа предназначена для работы в операционной среде Windows 7 и более поздних.

Данная программа представляет собой удобный и простой в использовании инструмент для наблюдения за положением искусственных спутников Земли (ИСЗ) в области видимости антенны, установленной на движущемся аппарате. На рис. 6 приведен графический интерфейс главного окна программы.

Главное окно программы содержит: рисунок с изображением поверхности Земли, кнопки для загрузки альманахов с набором сведений о текущем состоянии навигационных систем: GPS и ГЛОНАСС, кнопку загрузки траектории движения аппарата с многоэлементной антенной решеткой, элемент выбора положений антенны при перемещении по траектории движения.

Функционал разработанного продукта позволяет, используя данные альманахов эфемерид ГЛОНАСС и GPS, а также траектории движения аппарата с антенной (загружаемых из .csv файлов), отображать на карте полные группировки ИСЗ ГЛОНАСС и GPS, расположенные над поверхностью планеты с указанием их видимости для антенны рассматриваемого движущегося аппарата (в зависимости от положения в конкретной точке траектории движения – оранжевая линия) за счёт цветовой маркировки (зелёный – видимый, красный – невидимый, спутники GPS обозначаются квадратиками, ГЛОНАСС – кружками).



**Рис. 6. Графический интерфейс основного окна программы**

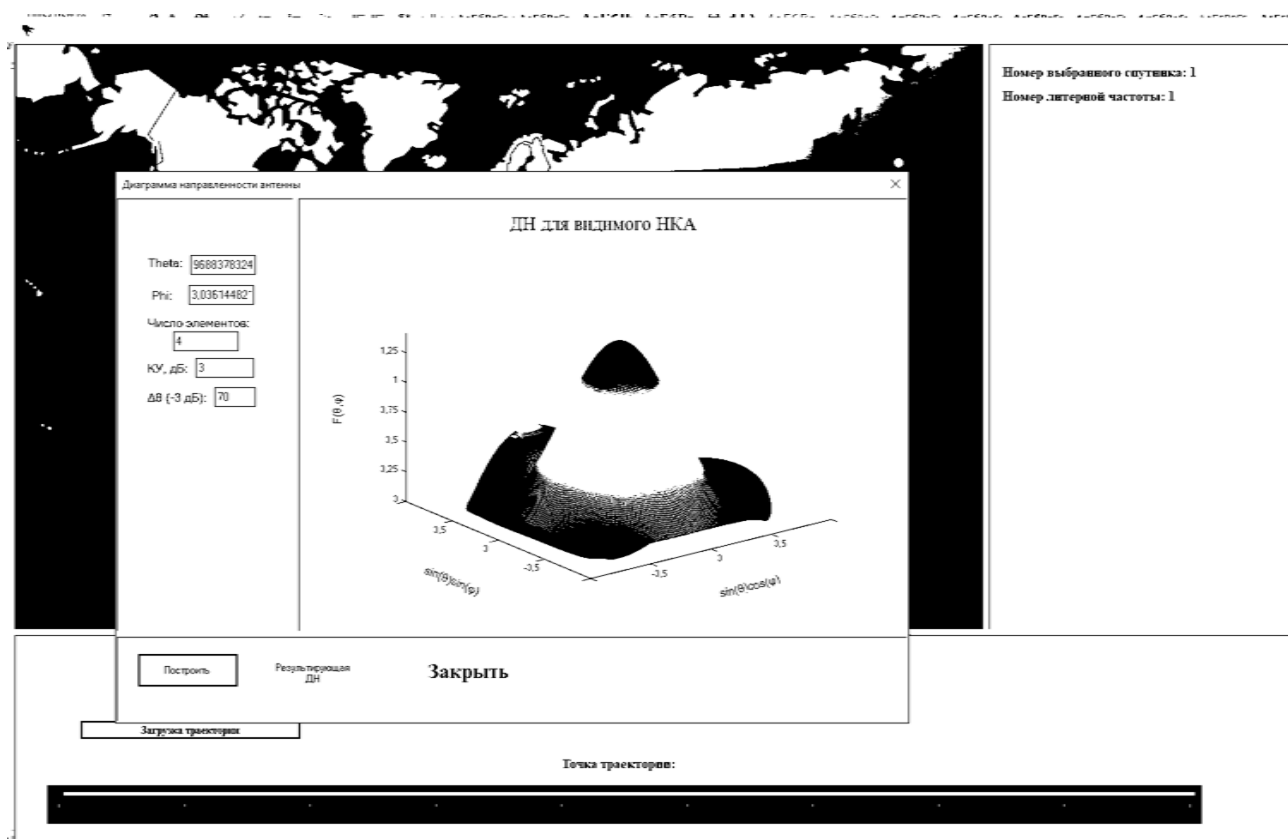
Загрузка альманахов с набором сведений о текущем состоянии навигационных систем осуществляется нажатием на кнопки «Загрузка данных GPS» или «Загрузка данных ГЛОНАСС». Если необходимо показать только одну из группировок спутников снимается флажок с элементов выбора «Спутники GPS» или «Спутники ГЛОНАСС», при снятии флажка соответствующая кнопка загрузки данных блокируется, и выводятся спутники только выбранной группировки.

Текущее положение антенной решетки на траектории определяется путем смещения бегунка элемента «Точка траектории».

В программе заложена функция построения трёхмерной диаграммы направленности антенной решётки, состоящей из изотропных излучателей, в направлении любого из спутников по выбору. Выбор спутника, на который направлен один из максимумов диаграммы направленности антенной решетки, осуществляется путем щелчка по видимому спутнику левой кнопкой мыши (для спутников ГЛОНАСС) и правой кнопкой мыши (для спутников GPS). Для выбранных спутников ГЛОНАСС справа от изображения Земли выводится номер спутника и номер литерной частоты, для спутников GPS только номер спутника.

На рис. 7 приведена 3D диаграмма направленности антенны, направленной на спутник ГЛОНАСС с номером 1.





**Рис. 7. Диаграмма направленности антенны на выбранный спутник**

Для построения диаграммы направленности антенной решетки нажимается кнопка «Построить», в окне «Число элементов» вводится количество элементов антенной решетки, в окне «КУ» – коэффициент усиления.

На рис. 8 изображена блок-схема программы. В качестве исходных данных используются альманахи расположения спутников ГЛОНАСС и GPS. Альманахи сохраняются в файлах с расширением .csv и загружаются в модуль расчета долготы и широты положения спутников. Далее эти данные поступают в блок обработки положения спутников, в котором с учетом траектории движения антенны и ее положения определяются видимые для антенны спутники. Видимые антенной спутники обозначаются зеленым цветом. При помощи щелчка мышью по нужному спутнику открывается окно с 3D-диаграммой направленности антенны.

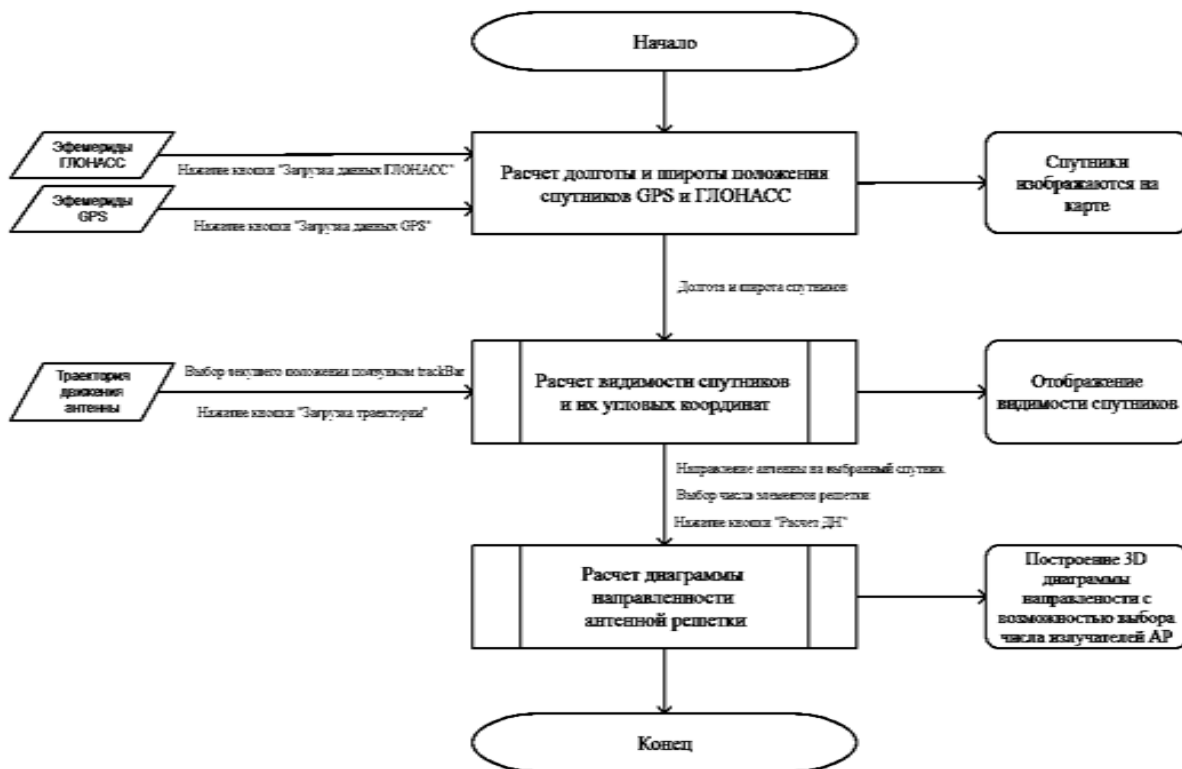


Рис. 8. Блок схема программы

### Заключение

Разработана математическая модель реализации пространственной селекции радиочастотных помех многоэлементной антенной решеткой аппаратуры спутниковой навигации путем формирования максимумов диаграммы направленности в направлении на навигационные космические аппараты; написана программа реализации алгоритма пространственной селекции помех на языке программирования высокого уровня C# в интегрированной среде разработки MS Visual Studio. Указанная программа для заданного момента времени обеспечивает определение доступной группировки НКА ГЛОНАСС и GPS в зависимости от текущего положения АСН и угловой ориентации АР, расчет и построение (визуализацию) диаграммы направленности многоэлементной антенной решетки для каждого видимого НКА, расчет и построение (визуализацию) результирующей диаграммы направленности многоэлементной антенной решетки. Приведено описание интерфейса разработанной программы и особенностей работы с ней.

### Библиографический список

1. **Гарин, Е.Н.** Измерение пространственной ориентации по сигналам спутниковых радионавигационных систем в условиях помех / Е.Н. Гарин, В.Н. Тяпкин, Ю.Л. Фатеев, Д.Д. Дмитриев, И.В. Тяпкин // *I-METHODS*. – 2020. – № 3. – с. 1-11.
2. **Ряполов, А.В.** Модель помехозащищенной антенной решетки для аппаратуры спутниковой навигации / А.В. Ряполов, Н.В. Фамбулов // *Вестник концерна ВКО «Алмаз-Антей»*. – 2018. – № 3. – с. 20-29.
3. **Chandran, S.** Adaptive Antenna Arrays. Trends and Applications / S. Chandran. – Berlin: Springer-Verlag, 2004. – p. 660.
4. **Ширман, Я.Д.** Радиоэлектронные системы: основы построения и теория. Справочник / Я.Д. Ширман, Ю.И. Лосев, Н.Н. Минервин, С.В. и др., под ред. Я.Д. Ширмана. Изд. 2-е, перераб. и доп. – М.: Радиотехника, 2007. – 512 с.
5. <https://www.glonass-iac.ru/>

6. ГЛОНАСС. Принципы построения и функционирования / Под ред. А.И. Перова, В.Н. Харисова. – 4-е изд., перераб. и доп. – М.: Радиотехника, 2010. – 800 с.
7. Глобальная навигационная спутниковая система ГЛОНАСС. Интерфейсный контрольный документ. Навигационный радиосигнал в диапазонах L1, L2 (редакция 5.1) – М.: Российский научно-исследовательский институт космического приборостроения, 2008. – 74 с.
8. **Гордеев, В.Н.** Кватернионы и трёхмерная геометрия / В.Н. Гордеев. – Киев, 2012. – 60 с.
9. **Григорьев, Л.Н.** Цифровое формирование диаграммы направленности в фазированных антенных решётках / Л.Н. Григорьев. – М.: Радиотехника, 2010. – 144 с.
10. **Баланис, К.** Введение в смарт-антенны / К.А. Баланис, И.П. Ионидес. – М.: Техносфера, 2012. – 200 с.
11. **Ишкова, Э.А.** С# Начала программирования / Э.А. Ивашкова. – СПб.: Наука и Техника, 2013. – 496 с.
12. **Шилдт, Г.** Полный справочник по С# / Г. Шилдт. — М.: Издательский дом «Вильямс», 2004. – 752 с.