

УДК 51-77

А.С. Ивлева, О.Е. Ошмарина, Е.В. Шадрина

АВТОМАТИЗАЦИЯ СБОРА И АНАЛИЗА ДАННЫХ В СОЦИАЛЬНЫХ СЕТЯХНациональный исследовательский университет Высшая Школа Экономики,
Нижний Новгород

Цель работы. Стремительно развивающиеся технологии поддерживают развитие онлайн социальных сетей. На данный момент все больше исследований посвящены взаимодействию между агентами социальных сетей. Методы программирования приложений (API методы) значительно упрощают проведение исследований в социальных сетях и позволяют автоматизировать сбор информации.

Научный подход. Для исследования было построено компьютерное приложение, которое собирает информацию о студентах трех крупных нижегородских вузов, а также представлен результат сетевого анализа.

Результат. В результате работы были исследованы возможности API разных онлайн социальных сетей; разработано приложение по сбору и анализу данных; представлены результаты сбора и общего анализа данных для 1500 пользователей.

Новизна. Возможность динамично изменять параметры разработанного приложения позволяет собрать информацию по совершенно иным характеристикам агентов, например место работы, хобби, музыка, использование других социальных сетей, занятие спортом, знание иностранных языков. Таким образом, появляется возможность провести иное социальное исследование.

Ключевые слова: api, онлайн социальная сеть, ВКонтакте, социальная структура, разработка приложения.

Введение

С развитием информационных технологий, поддерживающих удалённое взаимодействие, стремительно развиваются онлайн социальные сети. Ежегодно популярность и влияние онлайн социальных сетей растёт. Данный процесс связан с использованием виртуальных сетей, которые являются возможностью общения с людьми, удалёнными на тысячи километров. Взаимодействие с социальными сетями стало неотъемлемой частью жизни современного человека. Вследствие этого онлайн сети оказывают определенное воздействие на жизнь человека и его отношения с окружающими. Сейчас проводится все больше и больше исследований, связанных с созданием алгоритмов и приложений для изучения взаимоотношений между агентами социальных сетей.

Цель данного исследования – разработать приложение для автоматического сбора данных о пользователях онлайн социальной сети для последующего сетевого анализа.

В ходе работы были исследованы возможности API разных онлайн социальных сетей; разработано приложение по сбору и анализу данных; представлены результаты сбора и общего сетевого анализа данных для не менее 1500 пользователей.

Анализ социальных сетей - активно развивающееся направление западной и российской социологии. Исследования, направленные на изучение социальных сетей, используют различные подходы по оценки зрелости онлайн сообществ [9]. Также изучается вопрос автоматизации жизни современного человека при помощи онлайн социальных сетей. Рассматриваются два направления автоматизации - информатизация бизнеса и государственной деятельности и расширение спектра общения с друзьями, бывшими коллегами и однокашниками. Автоматизация позволяет более полно удовлетворять потребности в информации и в самовыражении, существенно увеличивает число инструментов самообслуживания [8, 10].

Интерес исследователей к этому направлению связан с тем, что оно предоставляет новый набор объяснительных моделей и аналитических инструментальных средств, которые находятся вне рамок обычных количественных методов. При этом в данной области накоплен богатый математический аппарат, позволяющий строить весьма сложные модели социальных взаимодействий, описывающие практически любые социальные системы [3, 5]. Социальные сети потенциально могут стать инструментом преобразования общества, распро-

странения научного и технического знания, формирования коллективов и общественных движений. В текущем положении они используются в основном для общения [9], маркетинговых исследований и продвижения продукции и услуг [4].

Интернет социальные сети также являются невероятно богатым источником данных о развивающихся социальных связях [7]. Экономическая теория формирования сети порождает гипотезы о том, какими правилами связей и экономических стимулов может быть сформирована архитектура сети, а также помогает определить их эффективность. Эти гипотезы можно эмпирически протестировать с данными из социальных сетей в Интернете [8].

Социальная сеть Facebook так же является обширной платформой для исследований [6]. Активно изучается влияние Facebook на формирование социальных групп среди студентов [3]. Также рассматриваются связи между социальной сетью и психологическим состоянием агента [5].

Помимо социальных сетей в настоящее время изучаются инструменты, с помощью которых можно работать с сетями и собирать информацию об агентах [7]. Прикладные API методы также активно используются в современных исследованиях. Опираясь на API методы разработчики могут создавать приложения, которые совместимы в контексте различных социальных сетей [4].

Анализ существующих платформ для разработки приложений онлайн социальных сетей Facebook и Вконтакте

В процессе исследования было необходимо создать приложение, которое бы собирало открытую информацию о пользователях из социальных сетей. Приложение может быть сделано по технологии Flash или IFrame [1]. Сайт Вконтакте изначально предоставляет разработчикам возможность работать с API методами. Наличие готовых функций значительно облегчает сбор данных и упрощает взаимодействие приложения с социальной сетью.

ВКонтакте и Facebook — основные онлайн социальные сети, на которые было ориентировано исследование.

С 2011 г. политика социальной сети Facebook требует наличия удаленного сервера для создания приложения. Руководство уточняет, что данное изменение не затронет работу популярных приложений, но создаст определенные препятствия для студентов и небольших групп [9]. Так как в процессе исследования не предполагалось создание собственного стороннего сервера, приложение для Facebook было не осуществимо в рамках настоящей работы.

Поэтому работа продолжалась только с социальной сетью ВКонтакте. Специфика ВКонтакте состоит в том, что в браузерах активной аудитории ресурса уже установлены последние версии Flash-плагинов, так как основные сервисы ВКонтакте (видеозаписи, аудиозаписи, загрузка фотографий, графики) уже построены на этой технологии. Технология Adobe Flash позволяет реализовать все, что может быть реализовано другими средствами, однако позволяет обойти кроссдоменные ограничения браузеров, а также решить проблемы безопасности старых версий Internet Explorer [2]. Поэтому для дальнейшей работы было выбрано Flash-приложение.

Программа Adobe Flash Professional CS5.5 является многофункциональной и позволяет разрабатывать Flash-приложения, поэтому для дальнейшей работы была выбрана именно она.

При создании приложения использованы стандартные API методы социальной интернет-сети ВКонтакте. Помимо них была задействована библиотека APIConnection, которая упрощает процесс запроса к методам.

Для обращения к ВКонтакте API необходимо использовать функцию объекта:

```
VK - VK.api(methodName: String, params: Object, [onComplete: Function, OnError: Function])
```

В случае удачного завершения запроса, функция `onComplete()` вернет объект, который содержит необходимые результаты. В случае неудачного завершения запроса, функция `OnError()` вернет объект, который содержит данные об ошибке.

При создании приложения были задействованы два API метода: `users.get()` и `friends.get()`. Параметры данных методов можно увидеть в табл. 1.

Таблица 1

Значения основных параметров для методов `users.get()` и `friends.get()`

Параметр	Описание
Uids	Для <code>users.get()</code> : Список id пользователей. Максимальное число – 1000. Для <code>friends.get()</code> : идентификатор пользователя, для которого необходимо получить список друзей. Если параметр не задан, он равен идентификатору текущего пользователя.
Fields: uid, first_name, last_name, nickname, sex, bdate, city, country, timezone, photo, photo_medium, photo_big, online, lists, contacts, education.	Поля анкет, которые необходимо получить.
name_case	Падеж для склонения имени и фамилии пользователя.
Count	Количество друзей, которое нужно вернуть (по умолчанию – все друзья).
Offset	смещение, необходимое для выборки определенного подмножества друзей.

Анализ полученных результатов

Интерфейс приложения и результат работы программы можно увидеть на рис. 1. Существует поле для ввода индивидуального номера пользователя ID. При нажатии кнопки “Get ID” в левом поле отобразятся индивидуальные номера всех друзей пользователя. При нажатии кнопки “Get INFO” в правом поле отобразится информация о друзьях пользователя.

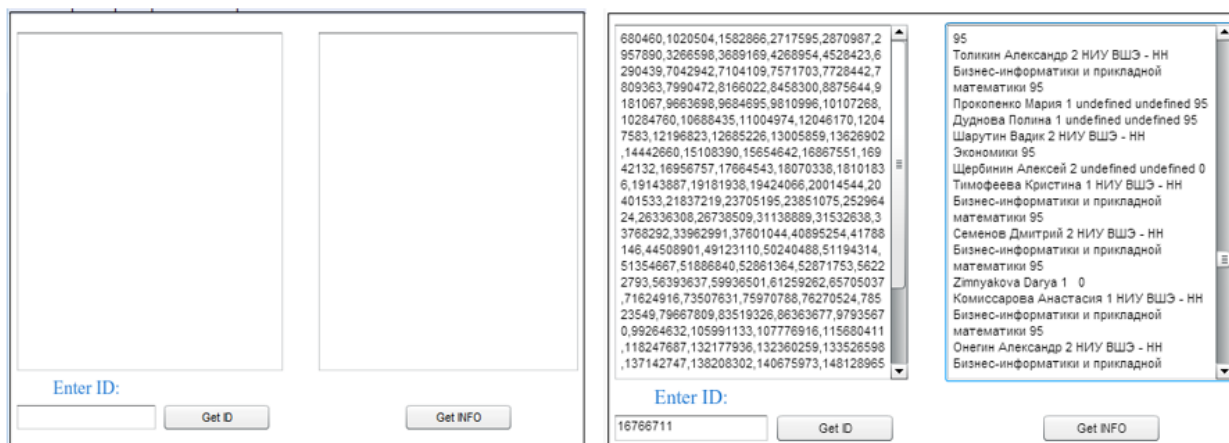


Рис. 1. Интерфейс приложения

В ходе работы с приложением были собраны данные о студентах в количестве 1500 человек с трёх учебных учреждений Нижнего Новгорода: НИУ ВШЭ – Нижний Новгород, ННГУ им. Н.И. Лобачевского, Нижегородская государственная медицинская академия (НижГМА).

В приложении были заданы параметры, которые собирали информацию об уникальном номере пользователя, имени и фамилии пользователя социальной сети, названии учебного учреждения и факультета, пол и город, в котором пользователь проживает. Все полученные данные экспортировались в Microsoft Office Excel.

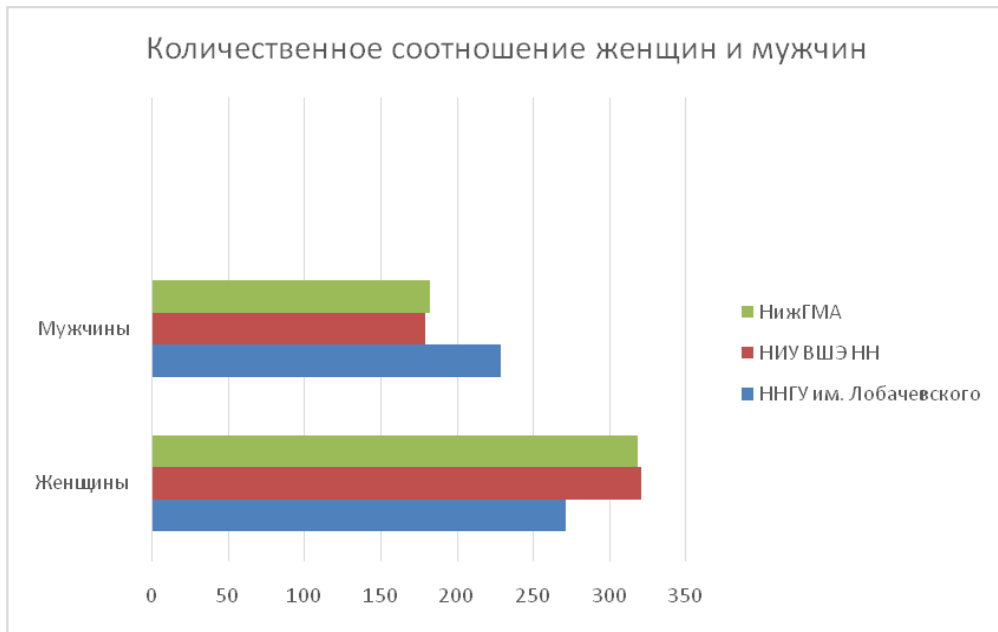


Рис. 2. Количественное соотношение женщин и мужчин

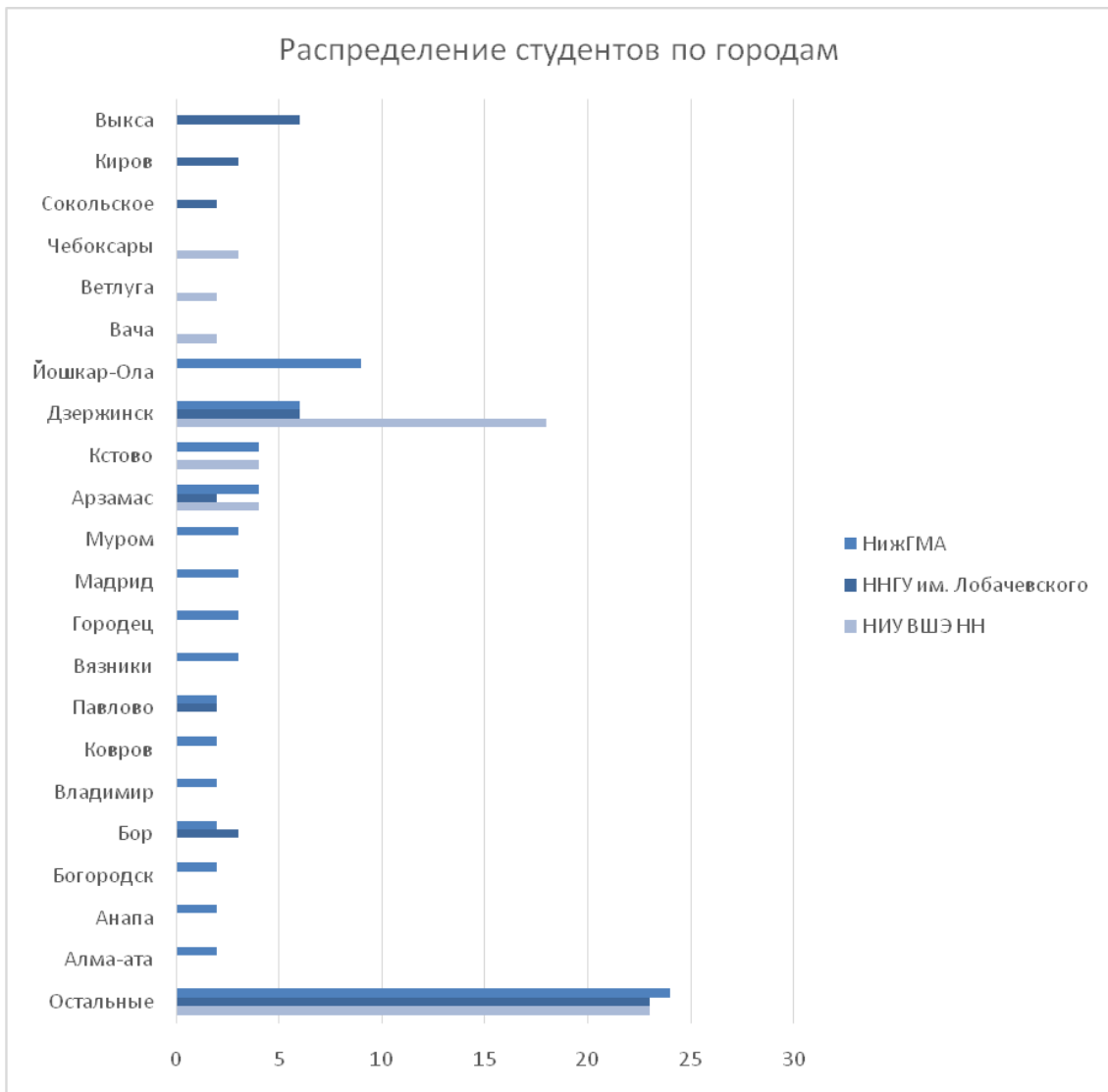


Рис. 3. Распределение по городам

В ходе сбора и обработки информации обнаружилось следующие погрешности: пользователь может указать в поле «город» место, где находится его учебное заведение, а не город, откуда пользователь приехал. В настоящей работе эти погрешности не устранялись, поскольку данные были собраны в тестовом режиме.

Также в ходе сбора информации были обнаружены некоторые особенности данных в онлайн-социальной сети: не все пользователи делают информацию на своей персональной странице доступной для просмотра, не все пользователи заполняют информацию о себе. Такие пользователи не включены в итоговую выборку.

Полученные данные были проанализированы по нескольким параметрам:

- количественное соотношение мужчин и женщин в каждом учебном заведении;
- количество приезжих студентов в каждом учебном заведении.

Как видно на диаграммах, представленных на рис. 2, девушек больше во всех трех учебных заведениях, причем в ННГУ им. Н.И. Лобачевского разница минимальна.

Данные позволяют рассмотреть распределение студентов по городам (рис. 3). Оказалось, что в ННГУ им. Н.И. Лобачевского в основном учатся из Выксы и Дзержинска. В НИУ ВШЭ – Нижний Новгород – из Дзержинска и Арзамаса. В НижГМА наибольшее число иногородних студентов и, что интуитивно понятно, большое число иностранных студентов. Наибольшее число студентов из Дзержинска и Йошкар-Олы.

Данные, собранные в процессе работы с приложением, можно представить в виде таблицы с булевыми значениями (рис. 4). Подобный формат удобно использовать далее для построения матриц и графов. Все это дает возможности для дальнейших исследований взаимоотношений агентов социальной сети и сбора статистических данных.

	Экономики	БИ	Права	Менежмента
Жен	94	55	48	88
Муж	53	54	22	35

Рис. 4. Соотношение мужчин и женщин по факультетам

Заключение

Данное исследование предполагало разработку приложения для автоматического сбора данных о пользователях онлайн социальных сетей с целью дальнейшего анализа структуры. В результате работы были исследованы возможности API разных онлайн социальных сетей; разработано приложение по сбору и анализу данных; представлены результаты сбора и общего анализа данных для 1500 пользователей.

Возможность динамично изменять параметры разработанного приложения позволяет собрать информацию о студентах других учебных заведений, собрать информацию по совершенно иным характеристикам агентов, например, место работы, хобби, музыка, использование других социальных сетей, занятие спортом, знание иностранных языков. Таким образом, появляется возможность провести иное социальное исследование.

Библиографический список

1. **Reimers, S.** Adobe Flash as a medium for online experimentation: A test of reaction time measurement capabilities / S. Reimers, N. Stewart // Behavior Research Methods. 2007. Т. 39. №. 3. С. 365–370. URL: <http://link.springer.com/article/10.3758/BF03193004>
2. Internet Exploer Standarts Support Documentation: [Электронный ресурс]. URL: [http://msdn.microsoft.com/en-us/library/ff410218\(v=vs.85\).aspx](http://msdn.microsoft.com/en-us/library/ff410218(v=vs.85).aspx)
3. **Greenhow, C.** Online social networks and learning //On the Horizon. 2011. Т. 19. №. 1. С. 4–12. URL: <http://www.igi-global.com/article/online-social-networking-learning/51563>
4. **Häsel, M.** Opensocial: an enabler for social applications on the web //Communications of the ACM. 2011. Т. 54. №. 1. С. 139–144. URL: <http://dl.acm.org/citation.cfm?id=1866765>

5. **Goyal, S.** Social networks on the Web //The Oxford Handbook of the Digital Economy. 2012. С. 434. URL: <http://www.oxfordhandbooks.com/view/10.1093/oxfordhb/9780195397840.001.0001/oxfordhb-9780195397840-e-16>
6. **Ellison, N. B.** The benefits of Facebook “friends:” Social capital and college students’ use of online social network sites/ N. B. Ellison, C. Steinfield, C. Lampe //J ournal of Computer-Mediated Communication. 2007. Т. 12. №. 4. С. 1143–1168. URL: http://www.academia.edu/4982860/Social_capital_the_benefit_of_Facebook_friends
7. **Ackland, R.** Using Facebook as a data source and platform for e-researching social networks // refereed paper presented at the Fourth International Conference on e-Social Science. 2008. С. 18–20. URL: <http://www.opensourceintelligence.eu/ric/doc/Using%20Facebook%20as%20a%20Data%20Source.pdf>
8. **Чураков, А. Н.** Анализ социальных сетей // Социологические исследования. 2001. №. 1. С. 109-121. URL: <http://www.twirpx.com/file/194725/>
9. **Краснов, Ф.** Измерение уровня зрелости профессионального сообщества / Ф. Краснов, Р. Яворский // Принята к публикации в журнале «Бизнес-информатика», 2013. URL: <http://bijournal.hse.ru/data/2013/06/17/1287071381/9.pdf>
10. **Славин, Б.Б.** Конвергенция архитектур социальной и корпоративной информационной среды человека // Бизнес-информатика 2012. № 2(20). С. 3–9. URL: <http://bijournal.hse.ru/data/2012/06/28/1255711514/1.pdf>

*Дата поступления
в редакцию 18.02.2015*

A.S. Ivleva, O.E. Oshmarina, E.V. Shadrina

AUTOMATE THE COLLECTION AND ANALYSIS OF DATA SOCIAL NETWORKS

National Research University Higher School of Economics, Nizhny Novgorod

Research focus: Rapidly developing technologies support the development of online social networks. At the moment, more and more studies are devoted to interaction between agents of social networks. Methods of application programming (API methods) greatly simplify the research in social networks and allow you to automate the collection of information.

Research methods: For the study was to construct a computer application that collects information about students three major universities in Nizhny Novgorod, as well as provides results of network analysis.

Results: As a result of API features were investigated various online social networks; application designed to collect and analyze data; the results of collection and analysis of data for the total 1500 users.

Novelty: Opportunity to dynamically change the parameters of the developed application allows you to gather information on the characteristics of completely different agents, such as jobs, hobbies, music, the use of other social networks, sports, knowledge of foreign languages. Thus, it is possible to carry out other social sciences.

Key words: api, online social network, VKontakte, social structure, application development.