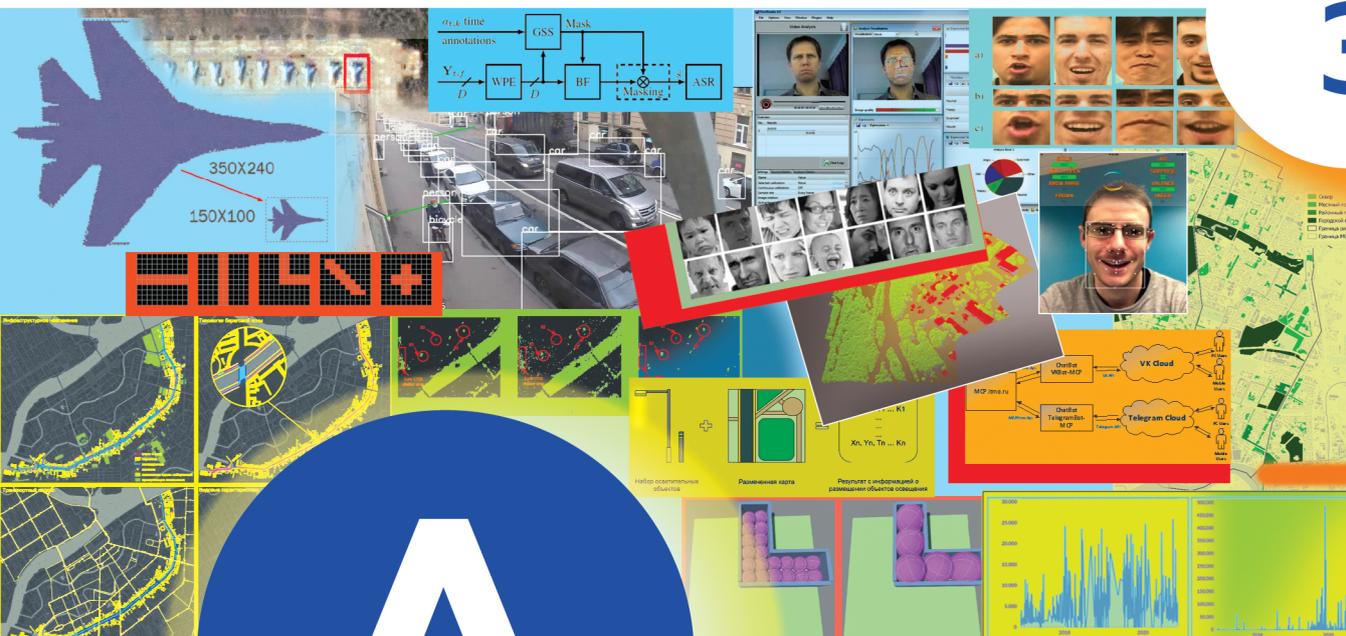




УНИВЕРСИТЕТ ИТМО

3



# АЛЬМАНАХ

НАУЧНЫХ РАБОТ  
МОЛОДЫХ  
УЧЕНЫХ

2021

**МИНИСТЕРСТВО НАУКИ И ВЫСШЕГО ОБРАЗОВАНИЯ  
РОССИЙСКОЙ ФЕДЕРАЦИИ**

**НАЦИОНАЛЬНЫЙ ИССЛЕДОВАТЕЛЬСКИЙ УНИВЕРСИТЕТ ИТМО  
(Университет ИТМО)**

**АЛЬМАНАХ  
НАУЧНЫХ РАБОТ  
МОЛОДЫХ УЧЁНЫХ  
Университета ИТМО**

Том 3



**УНИВЕРСИТЕТ ИТМО**

**Санкт-Петербург  
2021**

УДК\_082.2  
ББК\_94.3

**Альманах научных работ молодых ученых Университета ИТМО.** Том 3. СПб.:  
Университет ИТМО. 2021. 246 с.

Издание содержит результаты научных работ молодых ученых, доложенные на Пятидесятой научной и учебно-методической конференции Университета ИТМО по тематикам: информационные технологии и программирование; инфокоммуникационные технологии; дизайн и урбанистика.

#### РЕДАКЦИОННАЯ КОЛЛЕГИЯ

##### **Председатель редколлегии:**

*Бухановский Александр Валерьевич*

доктор технических наук, директор мегафакультета трансляционных информационных технологий Университета ИТМО.

##### **Члены редколлегии:**

*Митягин Сергей Александрович*

кандидат технических наук, директор института дизайна и урбанистики

*Басов Олег Олегович*

доктор технических наук, доцент, профессор факультета инфокоммуникационных технологий

*Матвеев Юрий Николаевич*

доктор технических наук, профессор факультета информационных технологий и программирования

*Краев Семён Владимирович*

кандидат филологических наук, декан факультета цифровых трансформаций

*Осипов Никита Алексеевич*

кандидат технических наук, доцент, доцент факультета инфокоммуникационных технологий

*Духанов Алексей Валентинович*

доктор технических наук, доцент, доцент факультета цифровых трансформаций

ISBN 978-5-7577-0648-1

ISBN 978-5-7577-0652-8 (Том 3)



Университет ИТМО (Санкт-Петербург) – национальный исследовательский университет, ведущий вуз России в области информационных, фотонных и биохимических технологий. Альма-матер победителей международных соревнований по программированию: ICPC (единственный в мире семикратный чемпион), Google Code Jam, Facebook Hacker Cup, Яндекс.Алгоритм, Russian Code Cup, Topcoder Open и др. Приоритетные направления: IT, фотоника, робототехника, квантовые коммуникации, трансляционная медицина, Life Sciences, Art&Science, Science Communication. Входит в ТОП-100 по направлению «Автоматизация и управление» Шанхайского предметного рейтинга (ARWU) и занимает 74 место в мире в британском предметном рейтинге QS по компьютерным наукам (Computer Science and Information Systems). С 2013 по 2020 гг. – лидер Проекта 5-100.

© Университет ИТМО, 2021

© Авторы, 2021

## **ВВЕДЕНИЕ**

Издание содержит результаты научных работ молодых ученых, доложенные 1-4 февраля 2021 года на Пятидесятой научной и учебно-методической конференции Университета ИТМО по тематике: информационные технологии и программирование; инфокоммуникационные технологии; дизайн и урбанистика.

Конференция проводится в целях ознакомления общественности с результатами научных исследований, выполненных в рамках: государственного задания Министерства науки и высшего образования РФ, стратегии развития Университета ИТМО до 2027 года, грантов Президента РФ для поддержки молодых российских ученых, грантов РФФИ, РНФ, по постановлению Правительства РФ N 218 от 9 апреля 2010 года " Об утверждении Правил предоставления субсидий на развитие кооперации российских образовательных организаций высшего образования, государственных научных учреждений и организаций реального сектора экономики в целях реализации комплексных проектов по созданию высокотехнологичных производств", по постановлению Правительства РФ № 220 от 09 апреля 2010 г. «О мерах по привлечению ведущих ученых в российские образовательные организации высшего образования, научные учреждения и государственные научные центры Российской Федерации», государственной поддержки центров Национальной технологической инициативы на базе образовательных организаций высшего образования и научных организаций, федерального проекта "Цифровые технологии", национальной программы «Цифровая экономика в Российской Федерации» и по инициативным научно-исследовательским проектам, выполняемыми преподавателями, научными сотрудниками, аспирантами, магистрантами и студентами Университета, в том числе в содружестве с предприятиями, организациями Российской Федерации, а также международными сообществами для увеличения эффективности научно-исследовательской деятельности и подготовки кадров и специалистов высшей квалификации.

**НАПРАВЛЕНИЕ  
«ИНФОРМАЦИОННЫЕ ТЕХНОЛОГИИ  
И ПРОГРАММИРОВАНИЕ; ИНФОКОММУНИКАЦИОННЫЕ  
ТЕХНОЛОГИИ; ДИЗАЙН И УРБАНИСТИКА»**



**Антонов Александр Сергеевич**  
Университет ИТМО,  
Институт дизайна и урбанистики,  
студент группы № С42551,  
направление подготовки: 09.04.03 – Прикладная информатика,  
e-mail: saha\_97@bk.ru



**Иванов Сергей Владимирович**  
Университет ИТМО,  
Институт дизайна и урбанистики,  
к.т.н., доцент,  
e-mail: svivanov@itmo.ru

УДК 004.8

**ИНТЕГРАЦИЯ НЕЙРОННОЙ СЕТИ  
ДЛЯ СЕМАНТИЧЕСКОЙ СЕГМЕНТАЦИИ  
ОБЪЕКТОВ ПО ОБЛАКУ ТОЧЕК В КОМПЛЕКСНУЮ  
СИСТЕМУ АНАЛИЗА И ОБРАБОТКИ ДАННЫХ  
А.С. Антонов, С.В. Иванов  
Научный руководитель – к.т.н, доцент С.В. Иванов**

Работа выполнена в рамках темы НИР «Интеграция нейронной сети для семантической сегментации объектов по облаку точек в комплексную систему анализа и обработки данных».

**Аннотация**

В работе представлены результаты подготовки данных для модели нейронной сети, а также была произведена интеграция модели для семантической сегментации облака точек в комплексную систему анализа и обработки данных. Были приведены результаты и анализ тестирования на реальных данных, определены перспективы дальнейших работ.

**Ключевые слова**

Семантическая сегментация, облако точек, машинное обучение, искусственный интеллект.

В компьютерном зрении [1] все более важной проблемой становится представление трехмерных сцен. Поэтому в последние годы облака точек становятся все более популярными для представления объемных данных.

Облака точек обычно используются для визуализации 3D-объектов, 3D-карт. Для различия объектов в облаке точек основным методом является семантическая сегментация (классификация объектов) [2]. Под семантической сегментацией понимается задача сегментирования изображений или облаков точек и определения того, что представляют собой различные сегменты.

В данной работе будет произведена интеграция нейронной сети для семантической сегментации объектов по облаку точек в комплексную систему анализа и обработки данных. Также необходимо подготовить данные облака точек для их использования в реализованной сети.

**Объект** исследования – информационные технологии.

**Предмет** исследования – методы глубокого обучения.

Для этого необходимо решить следующие задачи:

1. Подготовка данных аэросъемки для обучения нейронной сети.
2. Настройка сети под имеющиеся данные.
3. Визуализация полученных результатов.
4. Анализ полученных результатов.

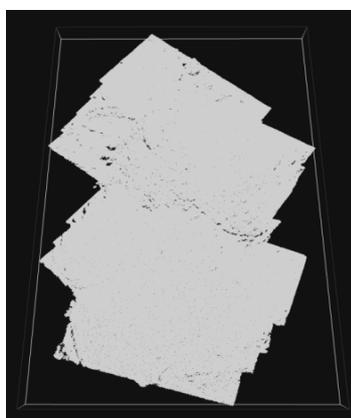
В данной работе будет производиться преобразование данных облака точек территории Магадана (рис. 1).



Рис. 1. Пример данных облака точек территории Магадана

Для разметки был использован инструмент Semantic Segmentation Editor [3]. Он представляет собой веб-инструмент маркировки для создания наборов данных для обучения искусственного интеллекта (2D и 3D). Инструмент был разработан в контексте исследования автономного вождения. Он поддерживает изображения (.jpg или .png) и облака точек (.pcd). Использует фреймворк Meteor и разработан на React, Paper.js и three.js.

Было произведено импортирование облака точек в Semantic Segmentation Editor (рис. 2).



а)



б)

Рис. 2. Импортированное облако точек без цвета (а) и с цветом (б)

Далее была произведена разметка данных облака точек. Происходит это посредством выделения объектов и отнесения их к какому-либо классу (земля, растительность, постройка и т.д.) (рис. 3).

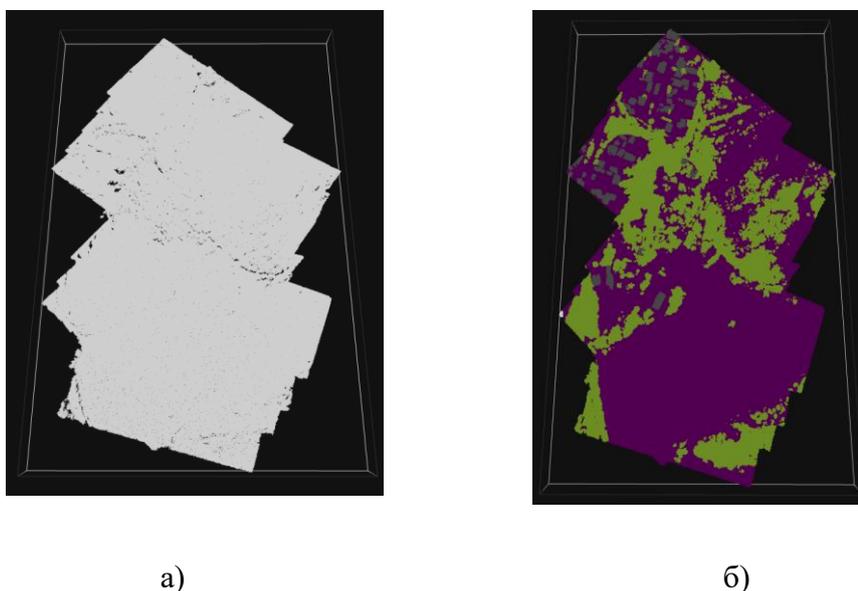


Рис. 3. Облако точек до разметки (а) и после (б)

После того, как вся интересующая нас область облака точек была размечена, необходимо преобразовать данные для дальнейшей подачи их на вход сети.

Для подачи данных на вход нейронной сети требуется поддержка одинакового формата. Поэтому был реализован скрипт со следующим алгоритмом:

1. Создается менее плотное облако точек, а также структурируются данные, путем вокселизации данных при помощи фильтра `voxel_downsize` библиотеки `pdal` [4]. За счет этого получаются облака точек с равным интервалом между точками.

2. Случайным образом удаляется несколько точек так, чтобы их количество стало кратным тому, сколько планируется подавать на вход нейронной сети, допустим 5120.

3. Полученное облако точек разделяется на прямоугольные области, содержащие точки в том количестве, которое будет подаваться на вход сети – 5120. Это достигается командой `split` библиотеки `pdal`.

Результат работы алгоритма после третьего этапа можно наблюдать на рис. 4.

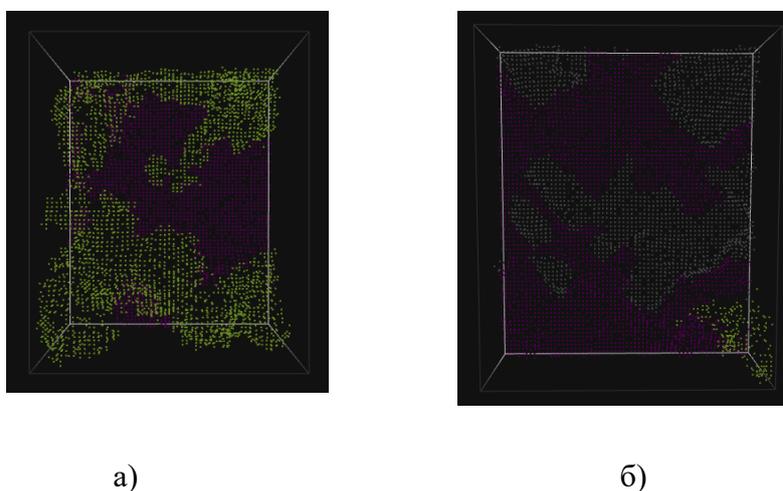


Рис. 4. Примеры облака точек после третьего этапа (а, б)

После подготовки данных для сети было произведено разделение данных на следующие выборки:

1. Train data (данные для обучения сети - 80%) (рис. 5).

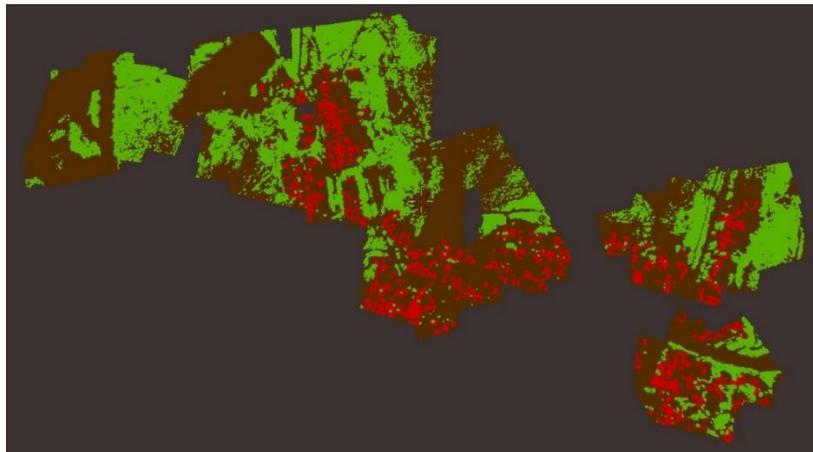


Рис. 5. Область обучающей выборки

2. Validation data (данные для проверки сети - 10%) (рис. 6).

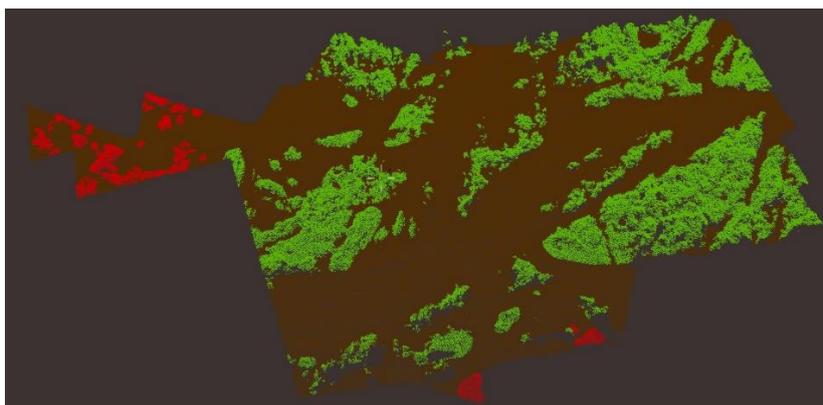


Рис. 6. Область валидационной выборки

3. Test data (данные для тестирования сети - 10%). (рис. 7)

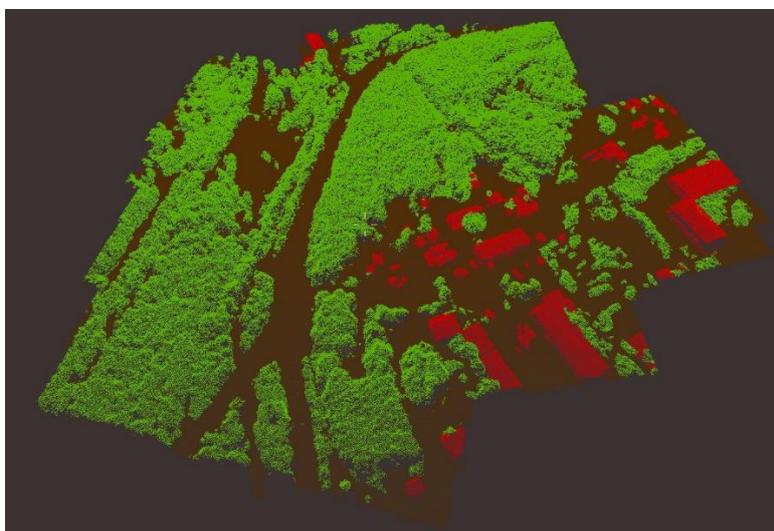


Рис. 7. Область тестовой выборки

Далее был произведен запуск сети. На вход подавались данные, где каждая точка имеет информацию о координате  $x, y, z$ , а также соответствующий класс.

Запустив сеть с параметром, задающим количество соседей точки для построения графа равным 10 ( $K = 10$ ), были получены следующие результаты:

$$\text{Accuracy} = 0.911, \text{IoU} = 0.778.$$

Далее была произведена настройка сети под имеющиеся данные. Помимо координат, данные точек имеют значение цвета в формате  $rgb$ . Можно использовать эти данные для улучшения распознавания сети. Необходимо расширить входной слой сети для возможности подачи туда данных о цвете точек.

Выполнив запуск сети со значениями  $rgb$  при  $K = 10$ , получаем следующие результаты:

$$\text{Accuracy} = 0.937, \text{IoU} = 0.826.$$

По результатам можно сказать, что наблюдается улучшение распознавания сети.

Также одним из настраиваемых параметров сети является значение « $K$ », которое указывает для точки, со сколькими ближайшими точками будет построен граф для дальнейших операций EdgeConv [5].

Было проведено аналогичное исследование показателей метрик для следующих значений  $K$ : 5, 10, 20, 30, и составлена сравнительная таблица.

Таблица

#### Показатели метрик с различными значениями параметра $K$

|          | $K = 5$ | $K = 10$ | $K = 20$ | $K = 30$ |
|----------|---------|----------|----------|----------|
| Accuracy | 0.933   | 0.937    | 0.940    | 0.913    |
| IoU      | 0.806   | 0.826    | 0.832    | 0.792    |

По результатам можно сказать, что при  $K = 20$  наблюдаются наилучшие показатели метрик accuracy и IoU.

Автоматическая разметка тестовых данных имеет вид (рис. 8).

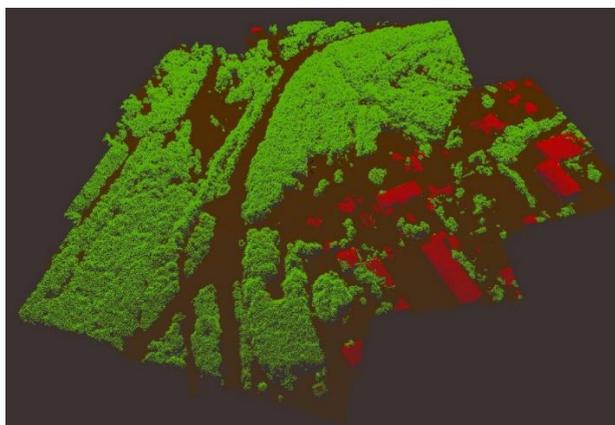


Рис. 8. Результат автоматической сегментации тестовой выборки

Имея итоговую модель нейронной сети, можно интегрировать ее в комплексную систему анализа и обработки данных.

В итоге последовательность действий алгоритма в комплексной системе анализа и обработке данных, связанная с облаком точек, является следующей:

1. При помощи определенного алгоритма происходит сшивка ортофотоплана из картинок, полученных при помощи беспилотника.
2. Создаются облака точек при помощи алгоритма построения плотных облаков из пар кадров.
3. Подготавливаются данные облака точек для подачи их на вход сети.
4. Выполняется применение модели нейронной сети.

В данной работе была произведена подготовка данных аэросъемки для обучения нейронной сети. Полученный датасет был протестирован на сети DGCNN [5]. Изначальный результат семантической сегментации составил 91.1%. После добавления дополнительных признаков в виде rgb цвета для каждой точки, а также настройки сети под имеющиеся данные наблюдается значительный прирост процента распознавания сети. Итоговый результат семантической сегментации для него составил 94%.

Далее в программе можно будет использовать полученное сегментированное облако точек в интересующих нас целях. Например, одной из актуальных задач может являться подсветка растительности, попадающей в охранную зону ЛЭП (линий электропередач), с высотой, превышающей норматив. Также можно подсвечивать здания или нефтепроводы в охранной зоне.

### Литература

1. Компьютерное зрение. Современный подход [Книга] / авт. Форсайт Д. [б.м.]: Диалектика-Вильямс. 2004.
2. Семантическая сегментация данных лазерного сканирования [Книга] / авт. А. Шаповалов Б. Р. Велижев В. [б.м.]: Москва. 2013.
3. Semantic Segmentation Editor [В Интернете] // github. 3 Сентябрь 2020 г. <https://github.com/Hitachi-Automotive-And-Industry-Lab/semanticsegmentation-editor>.
4. PDAL – Point Data Abstraction Library [В Интернете] // pdal. 19 Декабрь 2020 г. <https://pdal.io/index.html>.
5. Dynamic Graph CNN for Learning on Point Clouds [Отчет]/авт. Sun Yue Wang Yongbin. 2019.



**Антропов Даниил Валерьевич**

Год рождения: 1998

Университет ИТМО,

факультет инфокоммуникационных технологий,

студент группы № К4110с,

направление подготовки: 11.04.02 – Инфокоммуникационные

технологии и системы связи,

e-mail: avd980730@mail.ru



**Осипов Никита Алексеевич**

Год рождения: 1972

Университет ИТМО,

факультет инфокоммуникационных технологий,

к.т.н., доцент,

e-mail: nikita@ifmo.spb.ru

УДК 004.4

**ИССЛЕДОВАНИЕ ТРАНСПОРТНОГО  
ОБЕСПЕЧЕНИЯ И РЕШЕНИЕ ПРОБЛЕМ  
НА ОСНОВЕ ИНТЕЛЛЕКТУАЛЬНЫХ РЕШЕНИЙ**

**Д.В. Антропов**

**Научный руководитель – к.т.н. Н.А. Осипов**

**Аннотация**

В работе рассматривается применение интеллектуальной транспортной системы для устранения проблем в сфере дорожного движения. В качестве одного из решений рассматривается технология Vehicle-to-Everything, которая предполагает подключение транспортных средств друг к другу и к дорожной инфраструктуре. В ходе исследования делается вывод о существующих ограничениях при внедрении Vehicle-to-Everything технологии и предлагаются инструменты для решения некоторых из проблем.

**Ключевые слова**

Интеллектуальная транспортная система, Vehicle-to-Everything, Quality of Service, программно-конфигурируемые сети, периферийные сети.

Транспорт играет одну из главных ролей в обществе и экономике всего мира. Качество жизни в больших городах во многом зависит от безопасности, надежности и доступности транспортной системы. Но постоянный рост автопарка приводит к появлению следующих проблем:

- низкая пропускная способность дорог;
- высокий риск дорожно-транспортных происшествий;
- плохая организация общественного транспорта;
- нехватка парковочных мест.

Все перечисленные проблемы так или иначе связаны друг с другом, поэтому их решение требует комплексного подхода. Как можно понять, при хорошо спланированной и организованной транспортной инфраструктуре возрастает общая производительность города. Современным подходом для создания такой

инфраструктуры служит внедрение интеллектуальной транспортной системы (ИТС). ИТС отвечает за моделирование и регулирование транспортных потоков, сбор, анализ и предоставление информации конечному пользователю. ИТС призвана повышать безопасность и мобильность участников дорожного движения, а также автоматизировать операции, связанные с регулированием транспортных потоков. Все это достигается путем применения информационных технологий, которые способны в режиме реального времени прогнозировать и реагировать на ситуации на дороге.

Опыт зарубежных стран, где ИТС развивается намного лучше, чем в России, показывает, что внедрение интеллектуальных систем в отдельных случаях помогает достичь следующих результатов [1]:

- снижение дорожно-транспортных происшествий – до 60% на отдельных участках дороги;
- рост пропускной способности дороги – более 25%;
- оптимизация использования общественного транспорта – до 50%;
- увеличение пассажирооборота – на 20%;
- увеличение грузооборота в 1,5 – 2 раза.

Сейчас в сфере ИТС стало популярным рассмотрение возможности использования подключенных автомобилей. Область подключенных автомобилей относится к технологии V2X (Vehicle-to-Everything), которая объединяет в себе все связи между участниками дорожного движения и городской инфраструктурой: V2V (Vehicle-to-Vehicle), V2I (Vehicle-to-Infrastructure), V2P (Vehicle-to-Pedestrian), V2N (Vehicle-to-Network). На рисунке показана модель взаимодействия объектов транспортной среды с помощью технологии V2X [2].

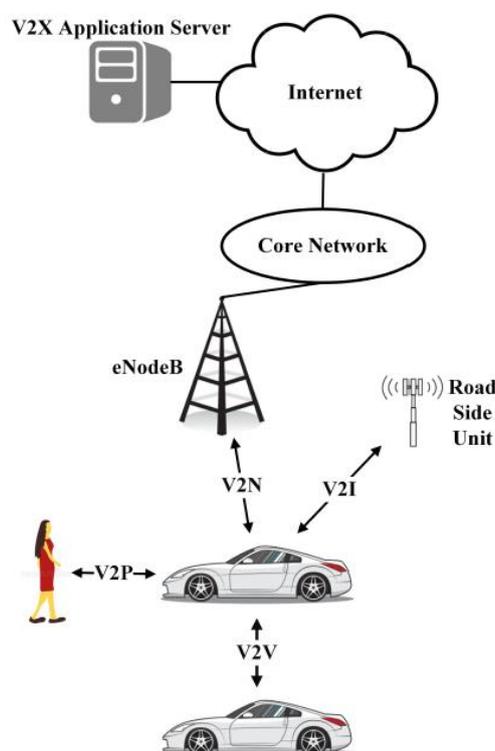


Рисунок. Модель взаимодействия внутри V2X сети

Предполагается, что транспортные средства, подключенные через V2X, смогут ориентироваться в дорожной обстановке, зная о положении других участников дорожного движения. Кроме того, транспортная инфраструктура также сможет

реагировать на дорожную ситуацию, что поможет сократить время в пути, сделает движение транспорта более предсказуемым и безопасным.

Примером использования V2X может быть светофор, который динамически адаптируется к транспортному потоку в разных направлениях и самостоятельно меняет режимы работы, чтобы избежать заторов и сохранить пропускную способность дороги на максимально возможном уровне.

Взаимодействие участников дорожного движения и инфраструктуры имеет много преимуществ в плане повышения безопасности, оптимизации транспортного потока и автоматизации транспортных процессов. Но внедрение технологии V2X требует решения определенных проблем, чтобы была возможность эффективно и действенно использовать эту систему.

Имеющийся набор проблем связан с тем, что сервисы, которые будут работать внутри V2X, сети имеют разную направленность и, соответственно, имеют разные требования к пропускной способности, задержке и частоте. Одни сервисы отвечают за организацию автономного движения, вторые – за организацию безопасности движения, третьи – за эффективность движения, четвертые – за информационно-развлекательную составляющую.

Ниже указаны проблемы, которые нужно решить для внедрения и эффективной работы V2X технологии [3]:

- QoS (Quality of Service);
- безопасность и конфиденциальность;
- масштабируемость и стоимость сети;
- стандартизация.

Первая проблема связана с предоставлением услуг QoS. Часть поддерживаемых V2X сервисов имеют строгие требования к качеству обслуживания, особенно с точки зрения задержки и пропускной способности. Например, сервисы с информационно-развлекательным контентом терпимы к задержкам, но требуют высокой скорости подключения. А у систем, отвечающих за автономность и безопасность движения, имеются совершенно противоположные требования. Составление соответствующих правил для выполнения этих требований и предоставления гарантий QoS – достаточно сложная задача, учитывая тот факт, что на качество обслуживания влияют многие внешние факторы.

Далее следует наиболее острая проблема современных сетей – это проблема безопасности. Из-за использования различных типов связи объекты V2X системы подвержены разным видам атак, которые могут привести к дорожно-транспортным происшествиям, получению персональных данных или угону автомобиля. Для предотвращения подобных случаев средства безопасности должны иметь возможность отслеживать злоумышленников.

Проблема масштабируемости и стоимости сети связана с тем, что увеличение количества подключенных транспортных средств (по прогнозам, к концу 2022 года мировой рынок подключенных автомобилей вырастет на 270% по сравнению с 2018 годом [4]) сопровождается увеличением стоимости инфраструктуры, необходимой для ее поддержания. Также само оборудование в автомобиле, требуемое для подключения к сети, является серьезным экономическим препятствием для автопроизводителей.

И последняя, но не менее важная проблема, стоящая перед V2X технологией – это проблема стандартизации. Пока нет определенного стандарта, который бы использовался в коммуникации V2X. Сейчас есть 2 стандарта, один, введенный в 2010 году – IEEE 802.11p, основан на использовании WLAN и уже используется в DSRC-радиосвязи ближнего действия, второй – новый стандарт C-V2X, основан на применении 5G соединения. Главное отличие DSRC от C-V2X с точки зрения коммерческого использования – в том, что DSRC уже используется на практике, то есть

уже протестирован, предсказуемо развивается и знаком производителям оборудования и провайдерам, что нельзя сказать о C-V2X, который имеет лучшие характеристики и потенциал.

В таблице представлены некоторые сравнительные характеристики IEEE 802.11p и C-V2X стандартов [5].

Таблица

**Сравнительные характеристики IEEE 802.11p и C-V2X**

|   | IEEE 802.11p                        | C-V2X (5G)  |
|---|-------------------------------------|---|
| Задержка                                | <100 мс (в среднем)                 | 0.5-10 мс (до 500 метров)<br>10-100мс (500-2000 метров) |
| Радиус действия                         | До 1 км                             | До 2 км   |
| Ширина канала                           | 10/20 МГц                           | 10/20/40/60/80/100/...МГц                               |
| Используемые методы мультиплексирования | TDM (временное мультиплексирование) | TDM и FDM (частотное мультиплексирование)               |

Одним из решений имеющихся проблем может служить использование программно-конфигурируемых сетей, представляющих собой сеть передачи данных, в которой уровень управления сетью реализуется программно и отделен от устройств передачи данных. Преимущество использования такого подхода в коммуникации V2X заключается в том, что разные объекты сети, такие как транспортные средства и элементы инфраструктуры, смогут действовать в роли сетевых коммутаторов. Это позволит использовать единый интерфейс для управления созданной автомобильной сетью.

Следующим инструментом, способным решить некоторые проблемы V2X сети, являются периферийные вычисления. Периферийные вычисления стремятся предложить распределенные вычисления и хранение данных на конечных устройствах, а не в централизованной вычислительной среде. Это позволяет добиться более низких задержек, что важно для работоспособности приложений, работающих в режиме реального времени.

Также в создании эффективной системы V2X могут помочь методы машинного обучения (МО) и анализа данных. Они способны извлекать нужную информацию из собранных данных и обучаться на их основе. Использовать алгоритмы машинного обучения можно для прогнозирования скорости и плотности транспортного потока. Также методы МО можно использовать для обеспечения безопасности на дорогах, например, определять угрозы на основе аномального поведения кого-то из участников движения.

### **Выводы**

В ходе работы были рассмотрены актуальные проблемы транспортных систем. В качестве инструмента для решения найденных проблем было рассмотрено такое направление интеллектуальной транспортной системы, как Vehicle-to-Everything (V2X). Сейчас область V2X активно развивается, так как она тесно связана с быстроразвивающимся Интернетом вещей и распространением 5G сетей. Поэтому в данной работе большая часть исследования была уделена направлению Vehicle-to-Everything.

В ходе исследования было получено представление о существующих на данный момент ограничениях для технологии V2X, которые связаны в первую очередь с высокими задержками и низкой пропускной способностью сетей. Появление и широкое распространение сетей пятого поколения может помочь обойти эти ограничения, но из-за новизны 5G технологии появляются новые проблемы в области V2X. Эти проблемы

связаны в первую очередь с безопасностью подключенных устройств к V2X сети, надежностью сети и возможностью ее масштабирования. В работе предложены возможные решения этих проблем, а именно: применение периферийных вычислений, алгоритмов машинного обучения и создание программно-конфигурируемых сетей.

### Литература

1. Галенко Л.А. Интеллектуальные транспортные системы – решение транспортных проблем / Николаева Р.В. // Техника и технология транспорта. 2017. № 3 (4). С.12.
2. Moubayed A. Softwarization, Virtualization & Machine Learning for Intelligent & Effective V2X Communications / Shami A. // IEEE Intelligent Transportation Systems Magazine. 2020.
3. Елькин Д.М. На пути к Интернету вещей в управлении транспортными потоками: обзор существующих методов управления дорожного движения / Вяткин В.В. // Известия ЮФУ. Технические науки. 2019. № 5 (207). С.100-113.
4. Connected Cars Report: 125 million vehicles by 2022, 5G Coming. [Электронный ресурс]. URL: <https://internetofbusiness.com/worldwide-connected-car-market-to-top-125-million-by-2022/> (дата обращения 12.12.2020).
5. ITS-G5 (стандарт связи для автомобилей). [Электронный ресурс]. URL: <https://www.tadviser.ru/index.php/> Статья: ITS-G5\_(стандарт\_связи\_для\_автомобилей) (дата обращения 04.12.2020).



**Безбородов Андрей Константинович**

Год рождения: 1998  
Университет ИТМО,  
факультет цифровых трансформаций,  
студент группы №J41322с,  
направление подготовки: 01.04.02 – Прикладная  
математика и информатика,  
e-mail: akbezborodov@itmo.ru



**Деева Ирина Юрьевна**

Год рождения: 1994  
Университет ИТМО,  
факультет цифровых трансформаций,  
аспирант,  
направление подготовки: 09.06.01 – Информатика  
и вычислительная техника,  
e-mail: ideeva@itmo.ru

УДК 004.942

## **АЛГОРИТМ ПОИСКА МЕСТОРОЖДЕНИЙ-АНАЛОГОВ НА ОСНОВЕ МЕТОДОВ МАШИННОГО ОБУЧЕНИЯ**

**А.К. Безбородов, И.Ю. Деева**

**Научный руководитель – к.т.н., доцент А.В. Калюжная**

### **Аннотация**

В работе разработан алгоритм поиска месторождений-аналогов на основе кластеризации методом  $k$ -средних. Дополнительно был реализован метод поиска пропущенных параметров месторождений в рамках кластеров с аналогами с помощью регрессионных моделей. Эксперименты были проведены для оценки влияния отсутствия различных параметров на качество дальнейших предсказаний.

### **Ключевые слова**

Нефтегазовые месторождения, месторождения-аналоги, поиск параметров, снижение размерности, кластеризация, геологические параметры.

Для эффективной оценки рисков в исследовании нефтегазовых месторождений необходимо как можно более точное моделирование их геологических параметров. Важнейшим инструментом, позволяющим произвести оценку параметров нефтегазовой залежи, является метод нахождения аналогов месторождений. На данный момент экспертами чаще всего используются месторождения, аналогичные по своим параметрам исследуемому, для анализа его признаков и понимания экономической жизнеспособности нефтегазовых проектов. После определения кластера, включающего в себя месторождения-аналоги, открывается большое число возможностей по моделированию и оценки исследуемого месторождения [1].

В общем смысле месторождение можно считать аналогом в случае, если оно совпадает или является схожим с рассматриваемым объектом по некоторым известным у обеих залежей параметрам. Такое сходство позволяет сделать вывод о том, что и многие остальные его характеристики будут аналогичны параметрам месторождения, в котором они неизвестны, что позволяет, например, произвести оценку пропусков в данных, полученных с месторождений-аналогов [1]. Несмотря на кажущуюся простоту

рассматриваемой задачи, на данный момент до сих пор не существует общепринятой стандартизации в задаче поиска аналогов месторождений.

С точки зрения времени и трудозатрат наиболее простым методом поиска аналогов месторождений является рассмотрение нефтегазовых залежей, находящихся рядом с исследуемой областью [2]. Такие подходы с сильным задействованием экспертных знаний позволяют оценить неизвестные параметры исследуемого месторождения и определить возможные аномальные значения в сжатые сроки, однако не гарантируют требуемую высокую точность получаемых результатов [2]. Исходя из анализа таких подходов, возникает необходимость в рассмотрении других методов, которые бы позволили рассмотреть процедуру поиска месторождений-аналогов с учетом возникающих сложных связей между параметрами при минимальном вовлечении экспертных знаний.

Применительно к задаче поиска аналогов месторождений в нефтегазовой отрасли возможно применение и подходов, основанных на различных технологиях машинного обучения. Данные методы позволяют получить оценку неизвестных параметров на основе имеющихся значений в базе данных, построить вероятностные модели, характеризующие сложные зависимости между признаками с целью получения кластера с аналогами исследуемой залежи.

В связи с тем, что получение набора аналогов месторождений по своей сути является задачей кластеризации в многомерном пространстве, появляется возможность рассмотреть различные алгоритмы кластеризации для решения поставленной задачи. В работах [3, 4] предлагается использовать иерархическую кластеризацию для решения поставленной задачи по поиску аналогов месторождений. После предобработки исходных данных выполняется статистическое обучение подготовленной модели для получения кластеров месторождений-аналогов [3]. Однако стоит отметить, что авторы предлагают предварительно использовать регрессионные модели для предсказания пропущенных значений на полной базе данных, что может существенным образом повлиять на качество последующей кластеризации.

Целесообразно сформулировать список требований к подходу, который способен решить поставленную задачу наиболее эффективным способом. Выбранный алгоритм должен работать как с непрерывными, так и категориальными данными нужной размерности, а также иметь численные метрики качества выделения кластеров с аналогами для оценки того, насколько правильно была построена модель. Далее, важно, чтобы используемый подход умел обрабатывать данные, содержащие пропуски, а также обладал допустимой вычислительной сложностью. Интерпретируемость полученной модели играет важную роль при обработке данных в случае, если модель используется не экспертом. Наконец, необходимо оценить вовлеченность экспертных знаний, которые в идеальном случае должны отсутствовать.

Исходя из предъявленных требований, был сделан выбор в пользу реализации с применением технологий машинного обучения, которые позволяют при средней вычислительной сложности и незначительном применении экспертных знаний получить хорошую интерпретируемость обученной модели, которая способна работать с пропусками.

Анализ и тестирование предлагаемого алгоритма производился на производственной базе данных, содержащей 442 месторождения без пропущенных параметров. Для итоговой реализации был сделан выбор в пользу 10 параметров, наиболее точно характеризующих внутренний состав нефтегазовых месторождений: Porosity, Permeability, Depth, Netpay, Gross, Lithology, Tectonic Regime, Hydrocarbon Type, Structural Setting, Period. Данные параметры являются как непрерывными, так и категориальными, что требует дополнительной обработки. Выбранные параметры позволяют достаточно точно описать любое месторождение из базы данных.

Принимая во внимание большой разброс в значениях параметров и их количество, было принято решение уменьшить размерность входных данных алгоритма по обнаружению аналогов. Для этой задачи на следующем этапе был использован метод главных компонент. Опытным путем было установлено, что наилучшие результаты последующей кластеризации и оценке пропущенных параметров получаются путем снижения размерности исходных данных с десяти до трех компонент. На рис. 1 изображен график рассеяния, на котором данные отображаются с тремя наиболее описательными главными компонентами. Для визуализации компоненты, соответствующие каждому из месторождений, были окрашены в различные цвета в зависимости от класса параметра Hydrocarbon Type.

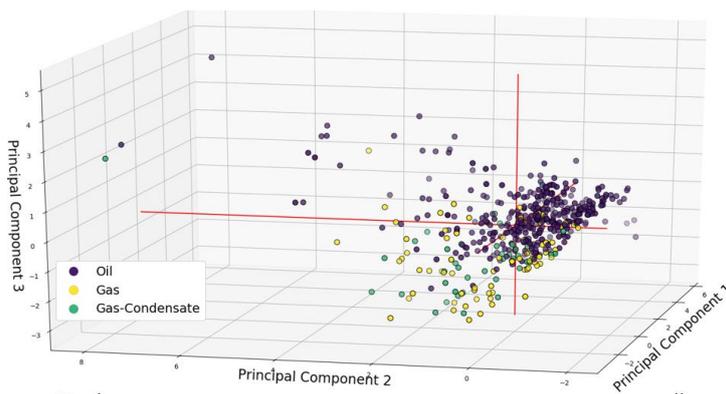


Рис. 1. График рассеяния для трех главных компонент, полученный в результате снижения размерности данных

Полученный рисунок демонстрирует достаточно четкое разделение между компонентой, характеризующей месторождения, для которых параметр Hydrocarbon Type принимает значение Oil, и компонентами, связанными с месторождениями, в которых значение параметра является Gas и Gas-Condensate.

После применения алгоритма по снижению размерности исходных данных возможно непосредственно применять различные алгоритмы кластеризации для формирования групп месторождений, обладающих схожими свойствами. Для данной задачи наилучшие результаты при проведении дальнейших предсказаний были получены с использованием кластеризации методом  $k$ -средних с числом кластеров, равным трем. Схематично примеры формирования кластеров с аналогами для кластеризации методом  $k$ -средних для некоторых непрерывных параметров представлены на рис. 2.

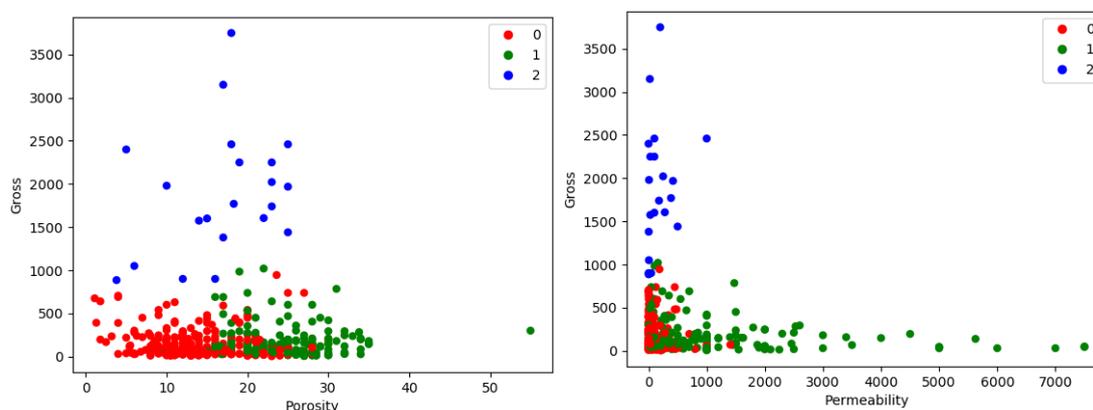


Рис. 2. Результаты формирования кластеров с аналогами для параметров Porosity и Permeability в зависимости от параметра Gross

Из рис. 2 видно, достаточно четкое разделение кластеров с месторождениями в зависимости от значений, принимаемых параметрами. Дополнительную оценку значений, принимаемых параметрами, в рамках различных кластеров можно произвести, построив распределения рассматриваемых параметров. Гистограммы таких распределений для параметров Depth и Gross приведены на рис. 3.

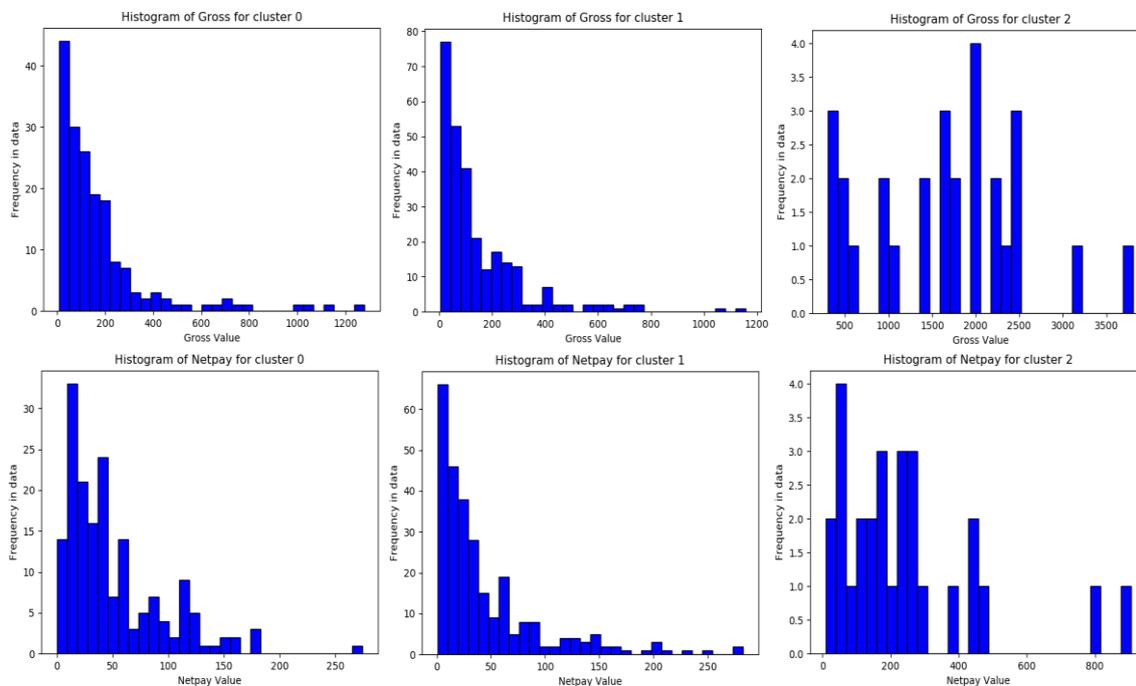


Рис. 3. Гистограммы распределений параметров для трех кластеров

Как видно из рис. 3, в результате к разным группам относятся данные, принимающие значения из различных интервалов, что говорит об эффективности проведенной кластеризации. Анализируя значения параметров месторождений, можно сделать вывод об их характере в рамках кластеров. К первому кластеру были отнесены 174 неглубоких нефтяных месторождений со средней пористостью и высокой проницаемостью. Второй кластер составил 276 нефтегазовых месторождений средней глубины с низкой пористостью и низкой проницаемостью. Наконец, к третьему кластеру, составившему 28 месторождений, были отнесены залежи, имеющие аномальные значения в параметрах Gross и Netpay.

Дополнительно оценивать эффективность проведенной кластеризации возможно с помощью карты, изображенной на рис. 4, на которой месторождения отмечены в зависимости от назначенного кластера.

Результатом вышеописанного алгоритма кластеризации является назначение кластера к каждому месторождению в исходной базе данных с месторождениями, у которых нет пропусков. Оценивать точность предсказаний пропущенных значений в рамках кластеров было решено с помощью регрессионных моделей, таких как многомерная линейная регрессия для непрерывных параметров и многомерная логистическая регрессия для категориальных. При этом с учетом вышеизложенного предсказания возможно выполнять различными способами:

- на всей базе данных (без кластеризации);
- в рамках кластера с учетом всех параметров;
- в рамках кластера с учетом параметров без пропусков.

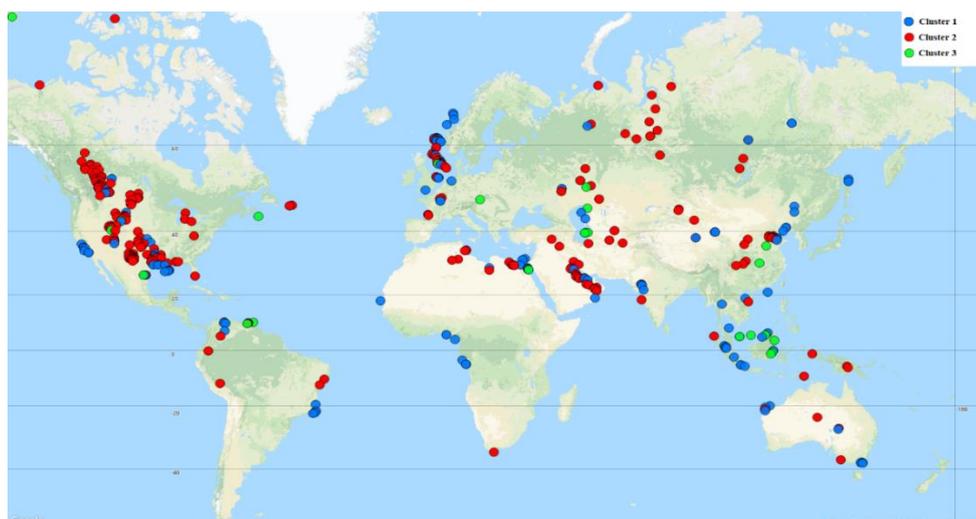


Рис. 4. Карта с месторождениями, окрашенными в результате кластеризации

Для выполнения тестирования в рамках каждого кластера месторождения были поделены на обучающую и тестовую выборки, составившие 75% и 25% соответственно. После оценки пропусков каждым из предложенным подходов формировалась средняя кластерная точность предсказания по метрике  $R^2$  для непрерывных параметров и средняя точность предсказания при оценке категориальных параметров. Полученные результаты приведены в табл. 1, 2.

Таблица 1

**Средняя точность предсказаний непрерывных параметров**

| Параметр     | $R^2$ Точность (на всей базе данных) | $R^2$ Точность (в рамках кластера с учетом всех параметров) | $R^2$ Точность (в рамках кластера без параметров с пропусками) |
|--------------|--------------------------------------|---|--|
| Porosity     | 0.2944                               | 0.4075  | 0.3546   |
| Depth        | 0.1501                               | 0.2834  | 0.2012   |
| Gross        | 0.1481                               | 0.2894  | 0.183733   |
| Netpay       | 0.2244                               | 0.4569  | 0.272233   |
| Permeability | 0.1535                               | 0.291   | 0.229033   |

Таблица 2

**Средняя точность предсказаний для категориальных параметров**

| Параметр           | Средняя точность (на всей базе данных) | Средняя точность (в рамках кластера с учетом всех параметров) | Средняя точность (в рамках кластера без параметров с пропусками) |
|--------------------|--|---|--|
| Lithology          | 0.5555                                 | 0.7266  | 0.642967   |
| Tectonic Regime    | 0.4861                                 | 0.5063  | 0.4595   |
| Hydrocarbon Type   | 0.7222                                 | 0.8111  | 0.8072   |
| Structural Setting | 0.2708                                 | 0.45  | 0.392333   |
| Period             | 0.2639                                 | 0.398   | 0.320933   |

Как видно из полученных таблиц, наилучшие результаты при предсказании пропущенных значений получены для месторождений, кластер с аналогами которых был известен заранее при обучении модели на всех параметрах. Результаты построения регрессионных моделей на основе полной базы данных без использования кластеризации оказались наихудшими. В случае некоторых параметров полученные результаты незначительно отличаются от модели, обученной на всех параметрах. Однако для некоторых параметров, таких как Gross и Netpay, разница в значениях достаточно оказалась достаточно существенной. Стоит отметить, что для многих параметров формирование кластеров с месторождениями-аналогами позволяет существенно улучшить точность оценки неизвестных значений.

В ходе работы на основе проведенного сравнительного анализа был реализован и протестирован алгоритм по поиску аналогов-месторождений с применением кластеризации методом k-средних и многомерной регрессии для предсказания пропущенных значений. Проведенный анализ позволил установить, что в результате применения рассмотренных алгоритмов кластеризации формируются группы месторождений со схожими исследуемыми параметрами. Средняя точность предсказаний для кластеров, не содержащих аномальные значения, оказалась примерно одинаковой (38% и 45% для непрерывных данных, 60% и 57% для категориальных). Отсутствие непрерывных параметров, таких как Netpay и Gross, пропуски в которых встречаются наиболее часто, приводит к заметному ухудшению в точности предсказаний.

В дальнейшей работе планируется исследовать, насколько отсутствие того или иного параметра влияет на точность предсказания пропусков в остальных параметрах. Также внимание стоит уделить и различным методам моделирования многомерных случайных величин, таким как байесовские сети, для решения поставленной задачи.

### Литература

1. Popova O. «Analogy in the World of Geological Uncertainties, or How Reservoir Analogs May Refine Your Probabilistic Geomodel» // SPE Annual Caspian Technical Conference and Exhibition. 2018. P. 1–13.
2. Voskresenskiy A. et al. «Variations in ranked list of reservoir analogs as an effect of search preferences» // Geosciences: Converting Knowledge into Resources. Saint-Petersburg, Russia. 2020. P. 1–4.
3. Martín Rodríguez H. et al. «New Approach to Identify Analogue Reservoirs» // SPE Economics and Management, New Orleans, Louisiana, USA. 2013. P. 1–17.
4. Da Gama e Silva R. et al. «Sensitivity analysis in a machine learning methodology for reservoir analogues» // Rio Oil & Gas Expo and Conference. Rio de Janeiro, Brazil. 2018. P. 1–9.



**Бейлин Максим Тимурович**

Год рождения: 1998

Университет ИТМО,

факультет инфокоммуникационных технологий,

студент группы №К41103с,

направление подготовки: 11.04.02 – Интеллектуальные

технологии в телекоммуникациях,

e-mail: icevenom98@gmail.com



**Осипов Никита Алексеевич**

Год рождения: 1972

Университет ИТМО,

факультет инфокоммуникационных технологий,

к.т.н., доцент,

e-mail: nikita@ifmo.spb.ru

УДК 004.6

**ОБЗОРНЫЙ АНАЛИЗ МЕТОДОВ ОБРАБОТКИ БОЛЬШИХ  
ДАННЫХ В УПРАВЛЯЮЩИХ СИСТЕМАХ ПРОГНОЗИРОВАНИЯ  
ПОВЕДЕНИЯ РЫНКА**

**М.Т. Бейлин**

**Научный руководитель – к.т.н., доцент Н.А. Осипов**

**Аннотация**

В работе рассмотрена проблема анализа поведения рынка ценных бумаг, акций, товаров и услуг. Для решения задач по обработке данных и принятия решений на их основе предложены методы анализа больших данных.

**Ключевые слова**

Информационно-управляющая система, рынок, большие данные, анализ данных, распределенные вычисления.

Темпы технического развития, развития экономики отдельных государств и человечества в целом приводят к появлению большого количества инструментов торговли, обмена информации, информационного влияния и воздействия на рынок. С течением времени сложность анализа поведения рынка возросла настолько, что не представляется возможной без применения вычислительной техники и специальных алгоритмов вычисления. Это также обусловлено как растущим количеством источников данных, так и их разрозненностью по структуре и содержанию. Для решения задач анализа поведения столь сложной структуры необходимо применять технологии анализа «Больших данных».

Необходимость анализа больших данных обусловлена стремлением перейти от количества к качеству, то есть извлечь полезную и практически применимую информацию из общего объема данных с целью создания новых продуктов и услуг, а также повышения качества управления экономическими процессами с целью увеличения доходности предприятий [1].

В анализе различных структур и процессов используются информационные управляющие системы (ИУС) – цифровые информационные системы управления некоторыми реальными объектами. Их основная функция – агрегация и обработка данных с целью построения решений по управлению объектами. Однако, несмотря на широкое распространение и стандартизацию подходов, применяемых в ИУС, их внедрение на практике оказывается дорогостоящим решением, так как требует тщательной настройки и поддержания на каждом конкретном предприятии [2].

С целью определить тип и формат данных, требующих обработки, был проведен анализ рынка ценных бумаг, товаров и услуг, а также исследований в данной области. Основными источниками данных являются открытые данные из регулярно публикуемых финансовых отчетов, аналитических прогнозов, пресс-релизов компаний и новостных статей. Также в качестве источников данных выступают числовые значения в виде цен или рыночных индексов. Дополнительной информацией о настроении и о покупательской активности населения могут служить данные из социальных сетей, данные поисковиков браузеров, а также данные GPS.

В существующих системах анализа используются такие методы, как:

- анализ трендов;
- анализ цикличности процессов, позволяющий предполагать на основе прошлых процессов варианты поведения рынка в будущем;
- методы экономико-математического моделирования, описывающие взаимосвязь рынка и влияющих на него факторов;
- пространственно-параметрическое моделирование – выявление основных факторов, формирующих и определяющих поведение рынка, и их последующий статистический анализ [3].

Приведенные подходы имеют различные реализации в существующих программных продуктах и сервисах, однако недостаточны для полноценного анализа и прогнозирования развития рынка, так как учитывают лишь малую часть определяющих факторов и зачастую применяются по отдельности. С целью усовершенствования имеющейся методологической базы и разработки подхода к анализу, учитывающего наибольшее количество факторов и источников данных, необходимо воспользоваться теорией больших данных и методов их анализа.

При изучении доступных алгоритмов анализа больших данных основной акцент ставился на такие показатели, как скорость обработки, обработка данных в режиме реального времени, а также возможность адаптации исходных данных. Основным фреймворком, используемым во всех алгоритмах распределенных вычислений, которые обеспечивают увеличение скорости обработки данных, является фреймворк MapReduce [4]. Суть MapReduce заключается в разбиении алгоритма анализа на этапы Map – обработка и фильтрация данных; Shuffle – разбиение на группы данных; Reduce – вычисление финальной функции для каждой группы данных.

Альтернативой фреймворку MapReduce является подход RDD – Resilient Distributed Dataset (устойчивый распространенный набор данных), который позволяет избежать необходимость разделения этапов Map и Reduce, а также обеспечивает отказоустойчивость за счет сохранения истории преобразования каждого экземпляра данных.

Были рассмотрены как классические алгоритмы анализа больших данных, наиболее известные из которых – Twister и Apache Spark, так и алгоритмы, построенные на теории графов, а именно Apache Giraph и GraphLab [5].

В таблице приведены особенности и преимущества рассмотренных подходов и технологий.

**Особенности алгоритмов анализа больших данных**

| № | Название      | Распределенные вычисления                                   | Масштабируемость   | Скорость обработки   | Требовательность к техническому оснащению                           |
|---|---------------|---|--|--|---|
| 1 | Map Reduce    | Разработан алгоритм, отвечающий за распределение вычислений | Механизмы перераспределения задач в зависимости от объема ресурсов | Зависит от размера сети, ограничена скоростью чтения и записи данных | Стабильная связь между узлами сети                                  |
| 2 | Hadoop        | Как у №1  | Как у №1   | Повышена за счет другого способа обработки данных                    | Как у №1  |
| 3 | Twister       | Как у №1  | Как у №1   | Повышена за счет обработки в памяти узлов                            | Повышенная отказоустойчивость из-за риска потери данных             |
| 4 | Apache Spark  | Как у №1  | Как у №1   | Выше за счет сокращения этапов анализа                               | Меньшие требования за счет RDD и последующего восстановления данных |
| 5 | Apache Giraph | Загрузка данных в узлы графов                               | Увеличение количества узлов графов пропорционально масштабу задачи | Выше за счет однократного распределения данных по узлам              | Отказоустойчивость узлов  |
| 6 | GraphLab      | Неоднородное распределение данных, общая память             | Как у №5   | Повышенная, но имеются риски простоя при сбое одного из узлов        | Скорость обмена данными между узлами и общей памятью                |

В рамках изучения ИУС прогнозирования поведения рынка наиболее важными показателями являются скорость анализа (большие массивы данных) и экономическая эффективность (анализ данных, позволяющий покрывать расходы извлекаемой потенциальной прибылью). Исходя из приведенных характеристик, наиболее применимыми являются Apache Spark и GraphLab.

Первая технология позволит анализировать большие массивы данных на предмет выявления закономерностей поведения рынка, при этом исключая большие траты на обеспечение высокого показателя отказоустойчивости. Вторая технология позволит анализировать зависимость рынка от сезона, состояния наполненности (конкуренции) рынка и других параметров, а также будет полезна при решении логистических задач за

счет применения теории графов. Оба подхода обеспечивают достаточно хорошие показатели отказоустойчивости, что позволит гарантированно получать достоверный анализ без потери данных.

Таким образом, были выбраны два алгоритма, наиболее полно отвечающих требованиям анализа данных при исследовании поведения рынка. В дальнейших исследованиях необходимо уделить внимание вопросу достоверности исходных данных и возможностям фильтрации, отсеивания и обнаружения аномалий и выбросов, которые могут негативно сказаться на качестве анализа. Также с целью выявления конкретных значений эффективности алгоритмов и их расширений требуется проведение тестового анализа данных с последующим сравнением и определением необходимых улучшений приведенных алгоритмов.

### Литература

1. Измалкова С.А., Головина Т.А. Использование глобальных технологий «BIG DATA» в управлении экономическими системами //Известия Тульского государственного университета. Экономические и юридические науки. 2015. №. 4-1.
2. Ковалева О.Н., Прохорова Н.А., Рыжова Е.С. Информационно-управляющие системы в экономической деятельности предприятий //Экономика и современный менеджмент: теория и практика. 2016. №. 3 (57). с. 26-34.
3. Абакумов Р.Г., Ходыкина И.В. Анализ существующих моделей для прогнозирования ценообразования на региональных рынках недвижимости //Инновационная экономика: перспективы развития и совершенствования. 2016. №. 1 (11).
4. Гладкий М.В. Модель распределенных вычислений MapReduce //Труды БГТУ. Серия 3: Физико-математические науки и информатика. 2016. №. 6 (188). с. 194-198.
5. Аксютин Е.М., Белов Ю.С. Обзор архитектур и методов машинного обучения для анализа больших данных //Электронный журнал: наука, техника и образование. 2016. №. 1. С. 134-141.



**Бокарева Алёна Александровна**  
Университет ИТМО,  
факультет инфокоммуникационных технологий,  
студент группы №К41403,  
направление подготовки: 09.04.03 – Прикладная информатика,  
e-mail: alena\_bokareva@itmo.ru



**Ватьян Александра Сергеевна**  
Университет ИТМО,  
факультет инфокоммуникационных технологий,  
к.т.н., ассистент,  
e-mail: aalexvatyan@gmail.com

УДК 004.622

## **СБОР И АНАЛИЗ ДАННЫХ ПО СЕРДЕЧНО-СОСУДИСТЫМ ЗАБОЛЕВАНИЯМ НА ФОНЕ COVID-19 ДЛЯ ВЫЯВЛЕНИЯ ЗАКОНОМЕРНОСТЕЙ И ВЕРОЯТНЫХ ПРОГНОЗОВ НА ОСНОВЕ СУЩЕСТВУЮЩЕЙ СТАТИСТИКИ**

**А.А. Бокарева**

**Научный руководитель – к.т.н. А.С. Ватьян**

Работа выполнена в рамках темы НИР №620178 «Методы и алгоритмы предсказания течения сердечно-сосудистых заболеваний путем комплексирования медицинской информации о сочетанных заболеваниях, в том числе COVID, с использованием нейронных сетей».

### **Аннотация**

В работе рассмотрены существующие методы сбора данных для анализа, а также анализ полученных данных для выявления закономерностей и вероятных прогнозов по сердечно-сосудистым заболеваниям (ССЗ) на фоне Covid-19 и, по итогам анализа имеющихся решений, предложен новый высокоэффективный метод, а также выявлен ряд гипотез по течению ССЗ на основании полученной выборки. Применение мобильного приложения «Единая персональная медицинская карта» позволяет решить не только задачи комфортного оперативного использования персональных медицинских данных, но и задачи аналитического построения прогнозов течения заболеваний (в том числе выявления предикторов развития ССЗ при Covid-19).

### **Ключевые слова**

Выявление предикторов, сердечно-сосудистые заболевания, Covid-19, мобильное приложение, использование медицинских данных, построение прогнозов течения заболевания.

Выявить закономерность, установить связи и принять превентивные меры для снижения, как задача минимум, летальных исходов по ССЗ на фоне Covid-19 и, как задача максимум, общего числа осложнений течения ССЗ при заражении данным вирусом является актуальной задачей не только с точки зрения медицины, но и для человечества в целом.

Ретроспективный анализ клинико-демографических показателей 1 007 больных COVID-19, госпитализированных в стационары (в отделение реанимации и интенсивной терапии, ОРИТ) Российской Федерации с острым респираторным дистресс-синдромом (ОРДС), обнаружил ССЗ у 61,4 % [1]. При сопутствующих ССЗ часто развиваются тяжелые формы COVID-19, требующие госпитализации больного в ОРИТ и ухудшающие его прогноз. В связи с этим у всех пациентов, в т. ч. со среднетяжелым и даже легким течением COVID-19, необходимо оценивать состояние ССС исходно и контролировать его в дальнейшем [2].

Сегодня имеются доказательства негативного влияния COVID-19 на развитие сердечно-сосудистой патологии *de novo* [3]. *N.S.Hendren et al.* предложено для обозначения кардиологических проявлений COVID-19 ввести новое понятие: острый COVID-19-ассоциированный сердечно-сосудистый синдром (*acute COVID19 cardiovascular syndrome, ACovCS*), описывающий широкий спектр сердечно-сосудистых и тромботических осложнений коронавирусной инфекции [4].

Однако, несмотря на доказательную базу и большое количество данных для прогнозирования течения ССЗ, на текущий момент выполнить поиск предикторов и установить четкие взаимосвязи не представляется возможным по причине разнородности данных, отсутствия четких прописанных протоколов исследований состояния здоровья у пациентов с ССЗ на фоне Covid-19, отсутствия единой базы данных по ССЗ и Covid-19.

Кардиальные проявления могут быть первичным феноменом при COVID-19 (по мнению ряда исследователей, это «сердечный фенотип» заболевания), но они также могут быть вторичными по отношению к легочному повреждению (смешанный легочно-сердечный фенотип) [5]. Важно отметить, что симптомы ССЗ встречаются в любой период госпитализации пациента, но, как правило, их риск возрастает с 15-го дня от начала лихорадки (или появления других симптомов вирусной инфекции). Развитие сердечно-сосудистых осложнений нередко развивается после стабилизации и / или улучшения респираторного статуса больного [5].

Были использованы и проанализированы следующие существующие способы сбора данных для решения этой задачи: деперсонализированные истории болезни пациентов стационаров с диагнозом Covid-19 и сопутствующими нарушениями со стороны ССС; статистические данные и результаты имеющихся исследований; анкетирование респондентов, перенесших Covid-19. У каждого из методов были выявлены ряд достоинств и недостатков: для сбора данных из историй болезни необходимо содействие и заинтересованность врачей (и, что немаловажно – наличие у них временных ресурсов на данное взаимодействие), однако данные, полученные таким образом, являются максимально информативными, не содержат субъективных искажений, но нуждаются в приведении в единый формат; для статистических данных (и данных из исследований) характерна избирательность, неполнота, их объективность также может вызывать некоторые сомнения; для данных, полученных методом анкетирования, характерными будут вариативность (как вследствие возможности корректировки опроса, так и благодаря когнитивным способностям респондентов), массовость, простота и доступность сбора. Однако для прогнозирования и выявления истинных закономерностей полученная выборка будет слишком мала и не может быть распространена на генеральную совокупность.

Был проведен анализ типовой истории болезни пациента, поступающего (и находящегося на лечении) в кардиологический стационар и в стационар с диагнозом Covid-19. По регламенту первичного осмотра и дальнейшего ведения (здесь имеются в виду обследования и анализы, проводимые в стационаре, без учета различий в терапии и протоколах лечения пациентов) на фоне Covid-19 существенных различий (кроме специфических маркеров на антитела к Covid-19 и ПЦР-диагностики) для людей вне группы риска не обнаружено. Для создания тестовой выборки нам необходимо определить из множества анализов и обследований те, которые будут максимально эффективно определять предикторы и создавать адекватную прогнозирующую базу. Такие параметры, как общие анализы крови и мочи, ряд биохимических показателей крови (печеночные, почечные, поджелудочные и другие), общий анамнез и аллергические реакции не представляют интереса в рамках данной работы и являются белым шумом.

Получение статистических отчетов оказалось сопряжено с рядом бюрократических сложностей, а информативность данных сводок и статистик столь невелика, что этот способ получения информации для построения модели машинного обучения не представляет собой ценности, и затраты на его реализацию превышают предполагаемую пользу. Использование данного ресурса возможно в качестве ознакомительного и направляющего. Анализ существующих на сегодняшний день медицинских статей на тему корреляции развития ССЗ и Covid-19 привел к однозначному пониманию, что взаимосвязь между возникновением и/или усилением проблем со стороны ССС и Covid-19 однозначная и крайне высокая. Анализ статей и учебных материалов показал, что «сбрасывать» из выборочных данных аспект медикаментозного лечения нельзя, поскольку ряд медикаментов, применяемых для лечения Covid-19, блокирует работу препаратов для поддержания функций ССС, и наоборот, часть препаратов, направленных на стабилизацию состояния со стороны ССС, препятствует полноценному лечению Covid-19.

Сбор деперсонализированных данных из медицинских учреждений на сегодняшний день ограничен ввиду наличия ограничений для работы в «красной зоне» и передачи каких-либо бумаг и электронных данных для работы с ними. Также ограничением служит отсутствие временных ресурсов у медицинских работников, осуществляющих непосредственные обязанности на местах. Однако данный вариант получения данных является наиболее емким, объективным и достоверным.

Опрос для перенесших заболевание Covid-19 был создан с использованием Гугл.Форм, выложен с доступом по ссылке и распространен с использованием различных социальных сетей и мессенджеров. Итоги анкетирования были сведены в общую таблицу, после чего обработаны и приведены в единый формат для удобства дальнейшей работы с данными опроса. Общее количество ответов по данному опросу составило 68 (в возрастной группе 35-50 лет, с уровнем дохода средний и выше среднего, перенесших Covid-19 в различной степени тяжести). При анализе полученной информации было выявлено следующее.

Для данной выборки характерно заболевание во «второй волне», в выборке оказались переболевшие в легкой форме (63 респондента – 94%) и средней степени тяжести (5 человек – 6%). При этом средняя продолжительность болезни составила 14 дней, характерных для большей части острых респираторных вирусных инфекций. Однако, наряду со стандартным течением и легкой формой заболевания, даже в представленной выборке появляются пациенты, болевшие более 3-х недель. Исходя из полученной информации на текущей выборке можно сделать вывод, что кризисными периодами при Covid-19 для возникновения осложнений и их развития являются 3 день, 7 день и 14 день заболевания, наиболее распространенным (в данной выборке) осложнением является возникновение тахикардии на 3-7-14 день заболевания. Сама по

себе тахикардия не является ССЗ, но служит предиктором для серьезных нарушений в работе ССС.

Проанализировав качество опросника и собранного материала, можно сделать вывод о необходимых доработках формы опроса и распространения среди потенциальных респондентов ссылки на него, а также о ключевых моментах, анализ и конкретизация которых позволит проводить машинное обучение и прогнозирование более детально.

Однако использование любого из трех способов для сбора информации показали себя малоэффективными в той или иной степени, не предоставляя максимально возможных данных для анализа.

Если существующие средства не дают искомый результат – необходимо создавать (адаптировать) новые. В связи с этим было принято решение о разработке (адаптации) специального программного обеспечения – «Единая медицинская персональная карта» для сбора необходимых деперсонализированных данных непосредственно от пациентов с документированного согласия запрашиваемых.

Разрабатываемое мобильное приложение представляет собой, с одной стороны, приложение, предназначенное для сбора с распределенных серверов разнообразной информации о конкретном лице, ее анализ, обработку и алгоритмы прогнозирования и планирования; с другой стороны – это интуитивно простой интерфейс для пользователя для получения, восприятия и использования полученных данных.

Конкурентные приложения не имеют алгоритмов для сбора данных с различных распределенных ресурсов и алгоритмов обработки данных разного формата для приведения их в единую форму для анализа и восприятия, что позволяет говорить о низком уровне конкуренции на текущий момент.

Для исследования и получения данных для прогнозирования течения ССЗ и Covid-19 возможность непосредственного получения необходимых данных в онлайн формате без когнитивных искажений является уникальной возможностью и создает перспективы полноценного анализа и прогнозирования.

Возможность получения всей информации о проведенных обследованиях, лабораторных исследованиях, диагностированных заболеваниях и сопутствующих патологиях в режиме онлайн доступа, без субъективных искажений и затрат дополнительных временных ресурсов со стороны медицинского персонала также позволяет собрать всю необходимую информацию для создания необходимой аналитической базы для построения прогнозирования течения ССЗ при Covid-19.

## Литература

1. Глыбочко П.В., Фомин В.В., Авдеев С.Н. и др. Клиническая характеристика 1007 больных тяжелой SARS-CoV-2 пневмонией, нуждавшихся в респираторной поддержке. Клиническая фармакология и терапия. 2020; 29 (2): 21–29. DOI: 10.32756/0869-5490-2020-2-21-29. / Glybochko P. V., Fomin V.V., Avdeev S.N. et al. [Clinical characteristics of 1007 intensive care unit patients with SARS-CoV-2 pneumonia]. Klinicheskaya farmakologiya i terapiya. 2020; 29 (2): 21–29. DOI: 10.32756/0869-5490-2020-2-21-29 (in Russian).
2. Бубнова М.Г., Аронов Д.М. COVID-19 и сердечно-сосудистые заболевания: от эпидемиологии до реабилитации. Пульмонология. 2020;30(5):688-699. <https://doi.org/10.18093/0869-0189-2020-30-5-688-699>.
3. Inciardi R.M., Adamo M., Lupi L. et al. Characteristics and outcomes of patients hospitalized for COVID-19 and cardiac disease in Northern Italy. Eur. Heart J. 2020; 41 (19): 1821– 1829. DOI: 10.1093/eurheartj/ehaa388.

4. Hendren N.S., Drazner M.H., Bozkurt B., Cooper L.T. Description and proposed management of the acute COVID-19 cardiovascular syndrome. *Circulation*. 2020; 141 (23): 1903–1914. DOI: 10.1161/CIRCULATIONAHA.120.047349.
5. Hendren N.S., Grodin J.L., Drazner M.H. Unique patterns of cardiovascular involvement in COVID-19. *J. Card. Fail.* 2020; 26 (6): 466–469. DOI: 10.1016/j.cardfail.2020.05.006.



**Брагина Татьяна Владимировна**

Год рождения: 1997  
Университет ИТМО,  
Институт дизайна и урбанистики,  
студент группы № С41801,  
направление подготовки: 27.04.07 – Научно-исследовательские  
технологии и экономика инноваций,  
e-mail: bragina.arch@yandex.ru



**Землянова Екатерина Игоревна**

Год рождения: 1995  
Университет ИТМО,  
Институт дизайна и урбанистики,  
студент группы № С4180,  
направление подготовки: 27.04.07 – Научно-исследовательские  
технологии и экономика инноваций.  
e-mail: katya-zemlyanova95@yandex.ru



**Лаушкина Анастасия Александровна**

Год рождения: 1996  
Университет ИТМО,  
Институт дизайна и урбанистики,  
студент группы № В42802,  
направление подготовки: 27.04.07 – Научно-исследовательские  
технологии и экономика инноваций,  
e-mail: nastasjalausckina@mail.ru



**Рослякова Светлана Витальевна**

Год рождения: 1988  
Университет ИТМО,  
Факультет инфокоммуникационных технологий,  
аспирант группы № 7772,  
направление подготовки: 09.06.01 – Информатика  
и вычислительная техника,  
e-mail: svetlana.roslyakova@itmo.ru



**Балаев Андрей Федорович**

Год рождения: 1981  
Университет ИТМО,  
Институт дизайна и урбанистики,  
к.т.н., доцент,  
e-mail: afbalaev@itmo.ru

УДК 004.89

## **ИССЛЕДОВАНИЕ ВЛИЯНИЯ РАЗЛИЧНЫХ СВЕТОВЫХ РЕЖИМОВ НА СНИЖЕНИЕ СТРЕССОВОГО СОСТОЯНИЯ ПОЛЬЗОВАТЕЛЯ**

**Т.В. Брагина, Е.И. Землянова, А.А. Лаушкина, С.В. Рослякова  
Научный руководитель – к.т.н., доцент А.Ф. Балаев**

Работа выполнена в рамках темы НИР №620185 «Разработка интеллектуальной многопользовательской адаптивной системы освещения для снижения конфликтного поведения и стрессовых ситуаций в образовательной среде».

### **Аннотация**

В работе рассмотрены механизмы влияния света на человека и его стрессовое состояние. Дана оценка негативным последствиям воздействия некачественно подобранных световых режимов. Основываясь на существующих исследованиях, были составлены основные режимы адаптивной системы освещения, применимые в офисных пространствах. Результатом стал подбор оптимальных характеристик светового оборудования для реализации проекта.

### **Ключевые слова**

Световые режимы, циркадные ритмы, характеристики света, спектральные характеристики, адаптивная система освещения, circadian lighting.

По статистике человек находится в офисном пространстве около 40 часов в неделю. Таким образом, большую часть светового дня мы проводим под искусственным освещением. При этом в течение дня человек может испытывать как всплески активности, так и усталость. Это происходит потому, что наши биологические часы и свет неразрывно связаны [1, 2]. Воздействие света на человека происходит через особые фоторецепторы в глазу, что, в свою очередь, оказывает влияние на наши циркадные ритмы [2]. Воздействие некачественного света приводит к негативным последствиям для физического и психического здоровья человека.

Целью данной работы являлось исследование различных световых режимов, позволяющих снизить стрессовое состояние пользователя и уровень конфликтного поведения в пространствах офисного типа.

Циркадным ритмом называют суточное изменение биологических процессов, протекающих в организме человека [1]. Такие ритмы находятся под контролем определённой области головного мозга, реагирующей на воздействия света благодаря нервным импульсам особых клеток сетчатки глаза. Сдвиг циркадных ритмов приводит к бессоннице, понижению работоспособности, тревоге и как следствие – стрессу. Данные сдвиги обусловлены изменением уровня гормонов: мелатонин отвечает за фазы сна, допамин – за настроение, кортизол – за активность, и т.д. Известно, что к невизуальным эффектам, влияющим на выработку гормонов, относятся следующие характеристики света:

- освещённость;
- спектральные характеристики;
- время воздействия;
- продолжительность воздействия.

Ранние отчеты исследований влияния света на человека и его стрессовое состояние предполагали, что люди в отличие от других организмов относительно нечувствительны к свету и более чувствительны к социальным сигналам, влияющим на

их психо-эмоциональное состояние [2]. Однако последующие эксперименты и наблюдения опровергли данное предположение, выявив фазозависимую реакцию на свет, складывающуюся от интенсивности света и от ответов на световые раздражители [1, 2].

Большинство исследований зависимости от интенсивности света обычно проводились путем применения световых стимулов одинакового спектрального состава в фиксированной продолжительности, одной циркадной фазе, но с изменением интенсивности света. Было обнаружено, что можно повлиять на психофизиологическое состояние человека при освещенности в пределах 100 люкс и ~ 1000 люкс [2]. При известных зависимостях влияния световых характеристик на человека мало изучен вопрос адаптации систем освещения к психо-эмоциональному состоянию пользователя. Естественная динамика изменений цветовой температуры света в природе положительно отражается на организме человека [4].

Согласно исследованиям, стрессовая реакция – это защитный механизм человека, помогающий ему справиться с трудностями, однако длительное и интенсивное проявление стресса причиняет вред здоровью [3]. Существует 3 стадии любой стрессовой реакции:

- стадия тревоги: запускается выброс гормонов адреналина и кортизола, увеличивается температура тела, повышается давление, ускоряется сердцебиение;
- стадия сопротивления: организм пытается сопротивляться и противодействовать изменениям в организме, которые произошли на предыдущей стадии;
- стадия истощения: организм теряет способность бороться со стрессом, длительность данного процесса приводит к истощению ресурсов организма, появляются отрицательные последствия для здоровья человека.

Если первая стадия помогает быстро находить решение стрессовых задач, то при переходе во вторую стадию организму следует помочь преодолеть данное состояние, чтобы не произошел эмоциональный перегруз организма. Вторая стадия – это лучший промежуток для влияния светом на снижение стрессового состояния при изменении освещенности и спектральных характеристик.

Взаимосвязь спектральных характеристик и циркадных ритмов привела к идее о том, что можно оказывать существенное влияние на периоды активности человека с помощью искусственного динамического освещения путем вариации цветовой температуры в течение рабочего дня [4]. Известно, что нейтральное освещение (4000 К) комфортно при условии решения стандартных задач, не обладающих эмоциональной окраской. Холодная температура (более 5000 К) повышает активность, способствует концентрации внимания, снижает сонливость. Данный тип светового режима повышает энергию человека, применяется в случаях сложных, быстрых задач, мозговых штурмов. Низкая цветовая температура (2800-3500 К) расслабляет организм, успокаивает возбужденное состояние, имеет особое влияние в случае постстрессовых ситуаций, восстанавливая ресурсы организма.

Основываясь на существующих исследованиях, был составлен основной набор режимов освещения, решающий проблему снижения стресса:

1. “Стандарт” 300 лк, 4000 К.
2. “Фокус на доске” 1000 лк, 4000 К и затемненное освещение аудитории 300 лк, 3800 К.
3. “Доска” 300 лк, 4000 К освещение в классе полностью выключено.
4. “Концентрат” 1060 лк, 5800 К.
5. “Активация” 675 лк, 6500 К.
6. “Расслабление” 325 лк, 3500 К.
7. “Максимальное расслабление” 275 лк, 3500 К.

Идея о разработке адаптивной системы освещения снижающей стрессовое состояние не новая. Смысл уже существующей концепции под названием Circadian Lighting сводится к тому, что активность человека значительно увеличивается при  $T_{цв} = 4000-6000$  К и снижается при  $T_{цв} = 2800-3500$  К [4]. Суточные ритмы могут сбиться, когда уровень вырабатываемого мелатонина может не достичь нормального показателя, поэтому работать в течение всего дня при цветовых температур  $\sim 5000-6000$  К вредно. Обратный эффект наблюдается при низких цветовых температурах, человек рано почувствует усталость, не будет продуктивен на работе.

Адаптивная система освещения должна быть способна динамически реагировать на текущие условия среды и потребности людей, учитывая их психо-эмоциональное состояние. Оптимальное световое оборудование для реализации имеет следующие показатели: светодиоды с настраиваемым спектром освещения. Допустимая средняя освещенность имеет разброс от 275 до 1060 ЛК. Световая температура от 3500 К до 6500 К. Индекс цветопередачи — от 80 до 90 (чем выше, тем лучше). В дальнейшем данная система позволит снижать стрессовое состояние пользователя, а также корректировать работоспособность и учитывать циркадные ритмы у человека.

### Литература

1. Munch M., Wirz-Justice A., Brown S., Kantermann T, Martin K., Stefani O., Vetter S. et al, «The Role of Daylight for Humans: Gaps in Current Knowledge» // *Clocks&Sleep* 2020. 2. P. 61–85.
2. Skene D., Duffy J., Czeisler C., «Effect of Light on Human Circadian Physiology» // *Sleep Med Clin*. 2009. 4(2). P. 165–177.
3. Шемякина О.О. Анализ причин стресса и методы его профилактики [Электронный ресурс] // *Психология и право*. 2012. Том 2. № 2. Режим доступа: <https://psyjournals.ru/psyandlaw/2012/n2/52061.shtml> (дата обращения: 18.12.2020)
4. Light and wellness: A circadian approach to lighting design [Электронный ресурс]. Режим доступа: <https://www.arup.com/perspectives/circadian-approach-to-lighting-design> (дата обращения: 18.12.2020).



**Бубнова Анна Викторовна**  
Университет ИТМО,  
факультет цифровых трансформаций,  
студент группы №7776,  
направление подготовки: 05.13.18 – Математическое  
моделирование, численные методы и комплексы программ,  
e-mail: avbubnova@itmo.ru



**Калюзная Анна Владимировна**  
Университет ИТМО,  
факультет цифровых трансформаций,  
к.т.н., доцент,  
e-mail: kalyuzhnaya.ann@gmail.com

УДК 004.942

## **МЕТОДЫ ПОИСКА НЕОДНОРОДНОСТЕЙ В СМЕШАННЫХ ДАННЫХ НА ОСНОВЕ БАЙЕСОВСКИХ СЕТЕЙ**

**А.В. Бубнова**

**Научный руководитель – к.т.н., доцент А.В. Калюзная**

Работа выполнена в рамках темы НИР №620175 «Виртуальная лаборатория разработки композитных моделей с помощью эволюционных методов автоматического машинного обучения».

### **Аннотация**

На практике основной интерес представляют данные смешанного типа, которые зачастую также содержат и пропущенные значения. Эта специфика не позволяет корректно использовать базовые методы поиска неоднородностей, построенные на сравнении интересующих нас объектов. В данной работе обсуждаются методы поиска неоднородностей, сосредоточенных вокруг фиксированного объекта, с использованием байесовских сетей для восстановления недостающих значений.

### **Ключевые слова**

Байесовские сети, поиск аналогов, восстановление пропусков, расстояния на разных типах данных, неоднородности.

С точки зрения статистики данные смешанного типа можно рассматривать как многомерную случайную величину – комбинацию категориальных и непрерывных переменных. Чаще всего речь идет о достаточно больших наборах данных, для которых ставится задача автоматического анализа без привлечения эксперта. К сожалению, стандартные и самые распространенные методы статистического анализа предполагают однородность, т. е. то, что представленная выборка взята из одной генеральной совокупности. Наличие неоднородностей сказывается на точности оценок, провоцирует падение качества моделирования, прогнозирования и восстановления. Это в свою очередь, приводит к дополнительным затратам на получение недостающей

информации и потерям при неверном планировании. Возникает вспомогательная задача проверки на однородность и поиска однородных фрагментах в неоднородных данных. Однако здесь есть несколько нюансов. Во-первых, часть методов применима только для одного типа данных, во-вторых, не все они могут работать с пропусками, в-третьих, не всегда ясно как оценивать однородность полученного фрагмента. Один из возможных подходов, позволяющий обойти эти ограничения, представлен в данной работе.

Поскольку задачи кластеризации, т. е. разбиения на подмножества из одной генеральной совокупности, зачастую NP-полны [1], здесь мы будем решать редуцированную задачу поиска однородного фрагмента, сконцентрированного вокруг некоторого фиксированного объекта. Рассмотрим этот объект и сравниваемые с ним как вектора в многомерном пространстве. Тогда для решения этой задачи можно использовать расстояния, например, косинусное расстояние для векторов  $u$  и  $v$ :

$$\cos(u, v) = 1 - \frac{u \cdot v}{\|u\|_2 \|v\|_2} = 1 - \frac{\sum_i u_i v_i}{\sqrt{\sum_k u_k^2} \sqrt{\sum_k v_k^2}}$$

Это расстояние имеет геометрический смысл: чем больше угол между векторами, тем ближе это расстояние к 1. Для совпадающих векторов, угол между которыми равен нулю, это расстояние тоже будет равно нулю. Аналогично можно определить взвешенное косинусное расстояние:

$$\cos_w(u, v) = 1 - \frac{\sum_i w_i u_i v_i}{\sqrt{\sum_k w_k u_k^2} \sqrt{\sum_k w_k v_k^2}}$$

в этом случае вводятся весовые коэффициенты, учитывающие важность определенных переменных. В данной работе будет рассмотрено невзвешенное косинусное расстояние, а также взвешенное, в котором веса выбираются таким образом, чтобы вклад каждой переменной на всем наборе относительно возможных целевых векторов в среднем был одинаковым.

Для лучшего понимания сформулируем последнее требование в виде системы уравнений. Для фиксированного вектора весов  $w = (w_1, \dots, w_p)$   $i$  – я переменная на полном наборе данных  $D$  относительно целевого вектора  $t$  дает следующий вклад:

$$\sum_{v \in D} \frac{w_i t_i v_i}{\sqrt{\sum_k w_k t_k^2} \sqrt{\sum_k w_k v_k^2}}$$

Тогда стоит задача подобрать такие веса  $w_i > 0$ , что для  $\forall i, j \in 1..p$ :

$$\sum_{t \in D} \sum_{v \in D} \frac{w_i t_i v_i}{\sqrt{\sum_k w_k t_k^2} \sqrt{\sum_k w_k v_k^2}} = \sum_{t \in D} \sum_{v \in D} \frac{w_j t_j v_j}{\sqrt{\sum_k w_k t_k^2} \sqrt{\sum_k w_k v_k^2}}$$

К сожалению, это не аналитическая задача, возможно, не всегда имеющая решение, так что вместо её решения мы ищем минимум следующей целевой функции:

$$\sum_{t \in D} \sum_{v \in D} \sum_{i, j \in \{1..p\}} \left( \frac{w_i t_i v_i}{\sqrt{\sum_k w_k t_k^2} \sqrt{\sum_k w_k v_k^2}} - \frac{w_j t_j v_j}{\sqrt{\sum_k w_k t_k^2} \sqrt{\sum_k w_k v_k^2}} \right)^2.$$

В данном случае рассматривается стандартный  $(p - 1)$ -симплекс весов:

$$\{w \in R^p: w_1 + \dots + w_p = 1, w_i \geq 0 \text{ for } i = 1, \dots, p\},$$

причем вектора с нулями игнорируют, потому что на соответствующих переменных мы бы потеряли необходимую нам информацию. Таким образом, на симплексе берется сетка с шагом  $1/m$ , а веса ограничиваются снизу фиксированным значением  $\varepsilon$ . Конкретные значения параметров и результат вычисления будут описаны далее.

Для того, чтобы обойти второе и третье ограничение, воспользуемся байесовскими сетями. С графической точки зрения байесовская сеть — это направленный ациклический граф (DAG). Он характеризует факторизацию совместного распределения, а также предоставляет информацию об условной независимости. Как граф его можно описать через совокупность вершин  $V = \{X_1, \dots, X_n\}$  и ребер  $E = \{(X_i, X_j)\}$ . Важным в контексте байесовских сетей является множество  $\Pi_{X_i}$  родительских вершин  $X_i$ , т. е. вершин, из которых в  $X_i$  проведено ребро. Под факторизацией, порожденной таким графом, подразумевается следующее:

$$P(X_1, \dots, X_n) = \prod_{i=1}^n P(X_i | \Pi_{X_i}).$$

Поскольку совместное распределение всегда можно факторизовать, оно представимо в виде некоторой байесовской сети. В данной работе мы будем использовать байесовскую сеть на полном наборе данных для восстановления пропущенных значений, а также сети на укороченном наборе для оценки однородности соответствующего фрагмента. Для обучения на данных используется МПТ [2] — относительно быстрый алгоритм, хотя и не самый точный.

Для фиксированного категориального значения целевого вектора считаем, что однородность выражается в частоте, с которой распределение на фрагменте принимает соответствующее значение. Чем чаще, тем ближе текущее распределение к вырожденному и однородному. Для непрерывных нас интересует отклонение наиболее выраженной моды распределения от фиксированного значения. Мода нас интересует, т. к. часто предположения о генеральной совокупности формулируются с помощью унимодальных распределений, и первый критерий однородности в таком случае — это проверка на единственность моды. Вторая причина заключается в том, что вокруг моды плотнее всего сконцентрированы значения и возрастает вероятность, что аппроксимация будет ближе к вырожденному распределению.

Перейдем к описанию эксперимента. Целью эксперимента являлась проверка однородности подвыборок, полученных с помощью косинусных расстояний, а также влияния предварительного восстановления пропущенных значений на однородность. Для анализа из набора на 500 объектов были выбраны 5 категориальных параметров: Tectonic regime, Hydrocarbon type, Period, Lithology, Structural setting. А также 5 непрерывных переменных: Gross, Netpay, Porosity, Permeability, Depth.

На данном наборе на сетке с шагом  $1/m = 1/10$ ,  $\varepsilon = 0.02$  с помощью ранее описанного метода для косинусного расстояния был найден вектор  $[0.02, 0.22, 0.22, 0.02, 0.22, 0.12, 0.12, 0.02, 0.02, 0.02]$ .

Опишем основные шаги эксперимента:

1. Берем объект из набора данных.
2. Фиксируем переменную и выбираем соответствующее значение в качестве целевого. Считаем, что это значение нам неизвестно.
3. Для методов с предварительным восстановлением пропусков обучаем байесовскую сеть на полном наборе данных без этого объекта и восстанавливаем пропущенное значение.

4. К данному объекту как к вектору подбираем 40 ближайших по метрике.
5. На этой подвыборке обучаем новую байесовскую сеть, которую и будем проверять на однородность относительно целевого вектора.
6. Сэмплируем с помощью байесовской сети из этого распределения в зависимости от типа переменной и смотрим частоту попадания в целевое значение или абсолютное отклонение моды.
7. Переходим к следующей переменной, а если они закончились, к следующему объекту.

Таким образом, когда для всех целевых векторов это будет рассчитано, можно для каждой переменной оценить ассурасу – частоту совпадения, а для непрерывных – MAE (mean absolute error):

$$MAE = \frac{\sum_{n=1}^N |\hat{y}_n - y_n|}{N}$$

В табл. 1 приведены результаты для категориальных переменных, а в табл. 2 для непрерывных.

Таблица 1

**Результаты эксперимента для категориальных**

| Метод<br>Параметр   | Cosine | Cosine +<br>восстановление | Cosine + веса | Cosine + веса +<br>восстановление |
|---------------------|--------|----------------------------|---------------|-----------------------------------|
| Structural setting  | 0.39   | 0.40                       | 0.36          | 0.38                              |
| Period              | 0.37   | 0.41                       | 0.34          | 0.46                              |
| Lithology           | 0.58   | 0.60                       | 0.56          | 0.59                              |
| Tectonic regime     | 0.31   | 0.37                       | 0.29          | 0.33                              |
| Hydrocarbon<br>type | 0.67   | 0.78                       | 0.67          | 0.77                              |

Таблица 2

**Результаты эксперимента для непрерывных**

| Метод<br>Параметр | Cosine  | Cosine +<br>восстановление | Cosine + веса | Cosine + веса +<br>восстановление |
|-------------------|---------|----------------------------|---------------|-----------------------------------|
| Depth             | 1004.62 | 887.43                     | 1097.53       | 932.74                            |
| Gross             | 437.40  | 440.44                     | 351.32        | 426.85                            |
| Netpay            | 80.12   | 92.13                      | 80.77         | 90.57                             |
| Porosity          | 6.98    | 7.17                       | 7.25          | 6.74                              |
| Permeability      | 1032.06 | 998.76                     | 997.89        | 902.97                            |

Результаты эксперимента показывают, что на категориальных переменных предварительное восстановление байесовскими сетями увеличивает однородность подборок, полученных с помощью рассмотренных метрик. А также то, что использование байесовских сетей позволяет для каждого метода наглядно оценить качество полученных фрагментов для каждой переменной.

В дальнейшем планируется переход к более точным смешанным байесовским сетям, сравнение результатов предварительного восстановления с помощью

байесовским сетей и существующих регрессионных моделей, рассмотрение комбинированных вариантов с восстановлением только категориальных переменных.

### Литература

1. Кельманов А.В. О сложности некоторых задач кластерного анализа // Журнал вычислительной математики и математической физики. Федеральное государственное бюджетное учреждение" Российская академия наук". 2011. Vol. 51, № 11. P. 2106–2112.
2. De Campos L.M. A scoring function for learning Bayesian networks based on mutual information and conditional independence tests // J. Mach. Learn. Res. 2006. Vol. 7, № 10. P. 2149–2187.



**Вобликов Станислав Дмитриевич**  
Университет ИТМО,  
факультет информационных технологий и программирования,  
студент группы №М41211,  
e-mail: vokilbov98@mail.ru



**Андрусенко Андрей Юрьевич**  
Университет ИТМО,  
научный сотрудник «ЦРТ-инновации»,  
e-mail: andrusenkoau@itmo.ru

УДК 004.08

## **ИЗУЧЕНИЕ РАЗЛИЧНЫХ РАСПИСАНИЙ ИЗМЕНЕНИЯ СКОРОСТИ ОБУЧЕНИЯ НЕЙРОННОЙ СЕТИ, ИСПОЛЬЗУЕМОЙ ДЛЯ АВТОМАТИЧЕСКОГО РАСПОЗНАВАНИЯ РЕЧИ**

**С.Д. Вобликов, А.Ю. Андрусенко**

### **Аннотация**

В работе рассмотрены различные расписания изменения скорости обучения глубокой нейронной сети, используемой в задаче распознавания речи, с целью повысить качество распознавания. Проведён эксперимент, показывающий влияние различных расписаний на качество распознавания.

### **Ключевые слова**

Нейронные сети, обучающие данные, скорость обучения, расписания, ассигасы.

Скорость обучения влияет на корректировку весов нейронной сети в зависимости от функции потерь. Чем ниже скорость, тем медленнее мы движемся к оптимальным значениям, но времени затратим больше. В связи с этим встает задача изучить влияния изменения скорости обучения во время тренировки нейронной сети, на качество распознавания.

Глубокая нейронная сеть состоит из четырёх скрытых слоёв, размерностью 1024 каждый. На вход подаётся вектор MFCC признаков с контекстом, равным пяти, на выходе вероятность для каждой из 51 фонемы.

Обучающая база состояла из десяти часов аудио данных. Обучение проводилось в десять эпох, размер батча 256. Тестовая база состояла из пятнадцати минут аудио данных [1, 2].

Были исследованы следующие расписания: LambdaLR, MultiplicativeLR, StepLR, MultiStepLR, ExponentialLR, CosineAnnealingLR, CyclicLR – triangular, CyclicLR - triangular2, CyclicLR - exp\_range, OneCycleLR – cos, OneCycleLR – linear, CosineAnnealingWarmRestarts.

LambdaLR. Скорость обучения на каждой итерации цикла умножается на некую функцию. В данном случае функция: номер эпохи  $** 0.65$  (рис. 1).

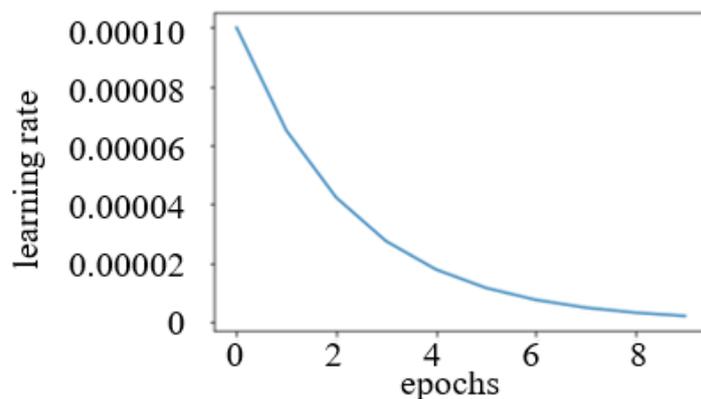


Рис. 1. Изменение скорости обучения по расписанию LambdaLR

MultiplicativeLR. Скорость обучения на каждой итерации цикла умножается на некий коэффициент. В данном случае коэффициент: 0.65 (рис. 2).

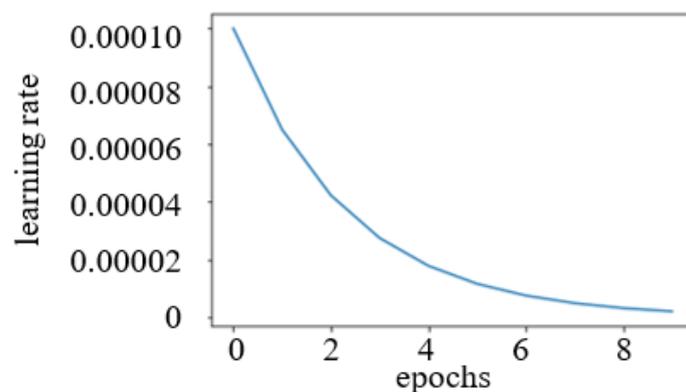


Рис. 2. Изменение скорости обучения по расписанию MultiplicativeLR

StepLR. Скорость обучения умножается на некий коэффициент через определённое количество итераций цикла. В данном случае коэффициент: 0.5, итераций 2 (рис. 3).

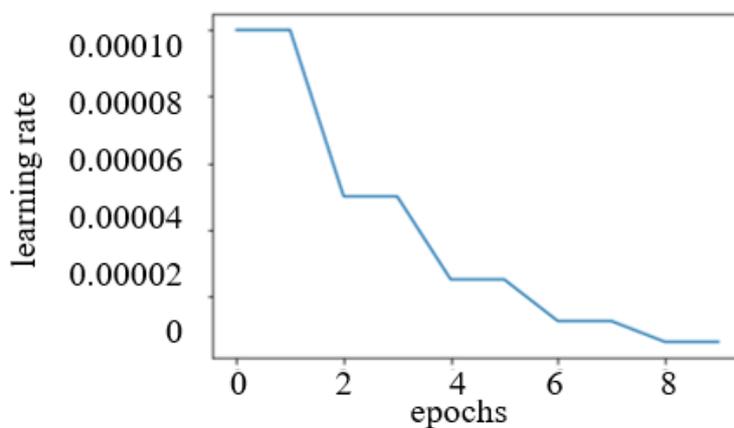


Рис. 3. Изменение скорости обучения по расписанию StepLR

MultiStepLR. Скорость обучения умножается на некий коэффициент на определённых итерациях цикла. В данном случае коэффициент: 0.1, итерации 3 и 8 (рис. 4).

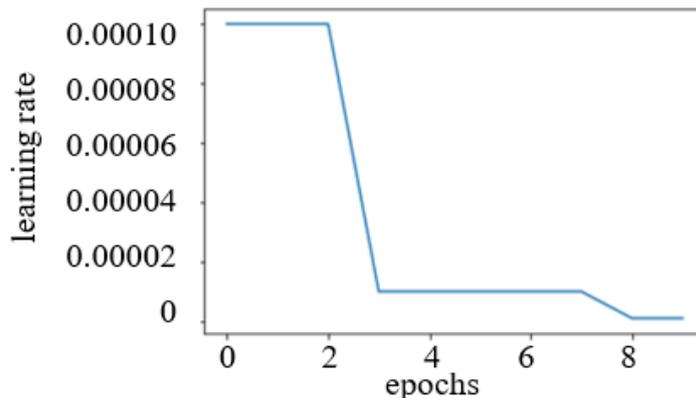


Рис. 4. Изменение скорости обучения по расписанию MultiStepLR

ExponentialLR. Скорость обучения на каждой итерации цикла уменьшается на заданный коэффициент. В данном случае коэффициент: 0.6 (рис. 5).

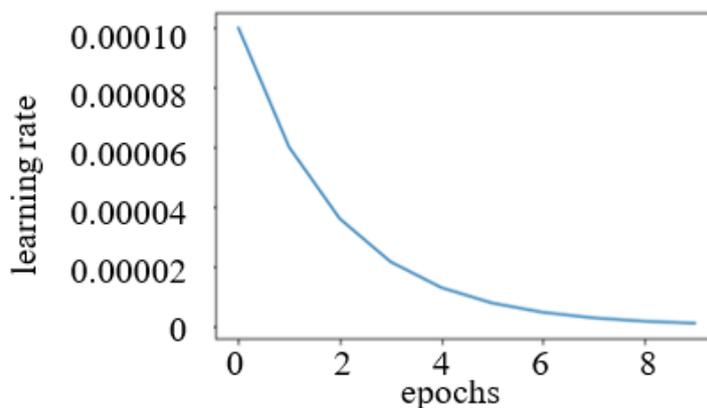


Рис. 5. Изменение скорости обучения по расписанию ExponentialLR

CosineAnnealingLR. Скорость обучения устанавливается с помощью косинусного графика отжига (рис. 6).

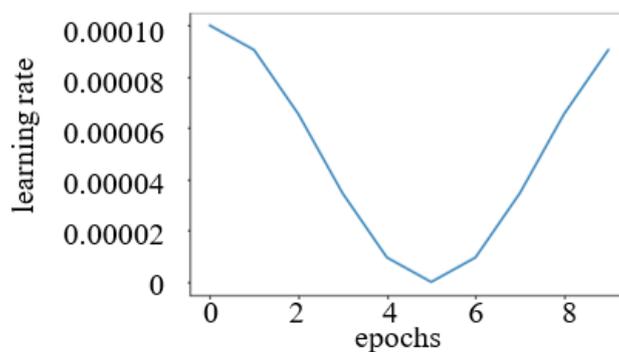


Рис. 6. Изменение скорости обучения по расписанию LambdaLR

CyclicLR – triangular. Скорость обучения перемещается между верхней и нижней границами в заданное количество итераций цикла. В данном случае: минимальное значение 0.0001, максимальное значение 0.01, итераций 2 (рис. 7).

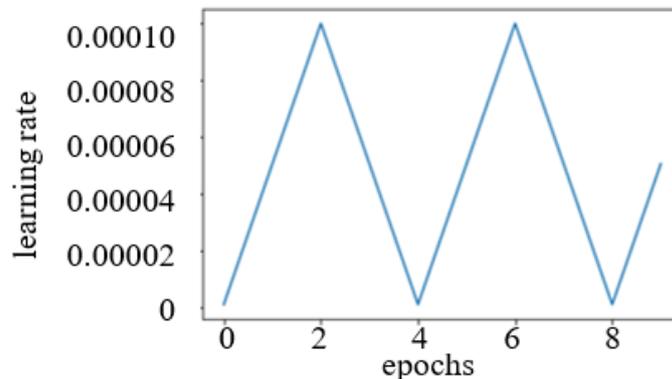


Рис. 7. Изменение скорости обучения по расписанию LambdaLR

CyclicLR - triangular2. Скорость обучения перемещается между верхней и нижней границами в заданное количество итераций цикла, верхняя граница уменьшается в два раза в заданное количество итераций. В данном случае: минимальное значение 0.0001, максимальное значение 0.01, итераций 2 (рис. 8).

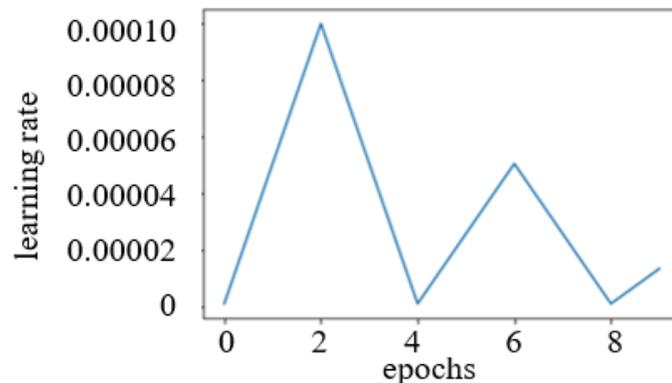


Рис. 8. Изменение скорости обучения по расписанию LambdaLR

CyclicLR - exp\_range. Скорость обучения перемещается между верхней и нижней границами в заданное количество итераций цикла, верхняя граница экспоненциально затухает в заданное количество итераций. В данном случае: минимальное значение 0.00001, максимальное значение 0.0001, итераций 2 (рис. 9).

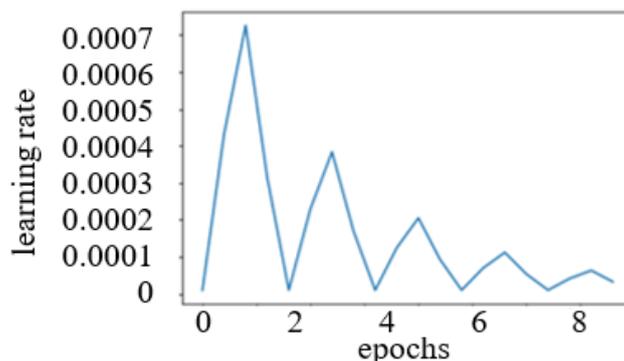


Рис. 9. Изменение скорости обучения по расписанию CosineAnnealingLR

OneCycleLR – cos. Скорость обучения увеличивается до заданного верхнего значения, после чего уменьшается. Верхнее значение 0.0001, начальная скорость обучения 0.00001 (рис. 10).

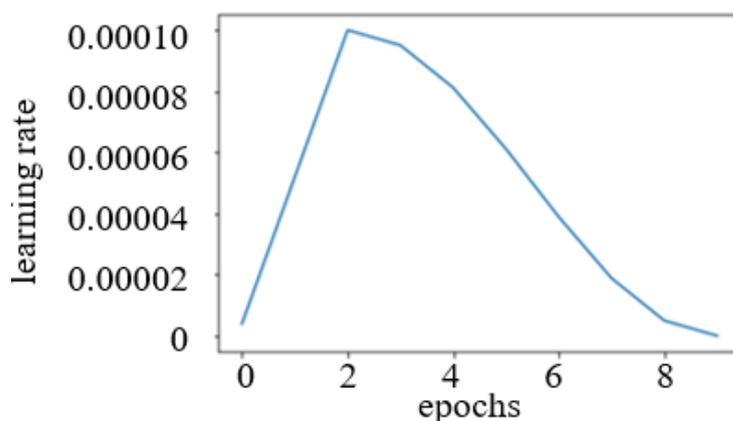


Рис. 10. Изменение скорости обучения по расписанию OneCycleLR – cos

OneCycleLR – linear. Скорость обучения линейно увеличивается до заданного верхнего значения, после чего уменьшается. Верхнее значение 0.0001, начальная скорость обучения 0.00001 (рис. 11).

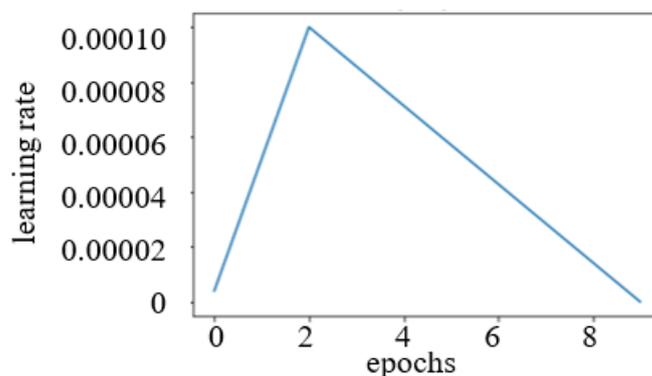


Рис. 11. Изменение скорости обучения по расписанию OneCycleLR – linear

CosineAnnealingWarmRestarts. Скорость обучения устанавливается с помощью косинусного отжига, после заданного количество итераций цикла, перезапускается. Количество итераций 2 (рис. 12).

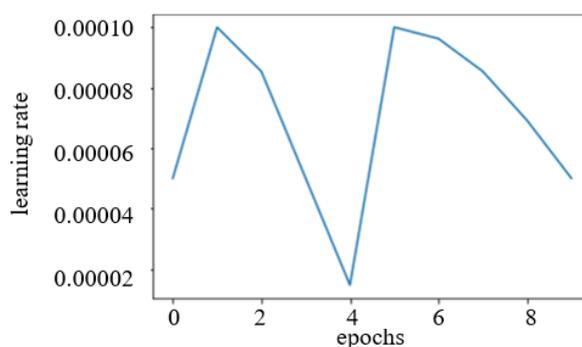


Рис. 12. Изменение скорости обучения по расписанию CosineAnnealingWarmRestarts

### Результаты эксперимента

| Scheduler         | Конечный loss | Frame accuracy | Scheduler              | Конечный loss | Frame accuracy |
|-------------------|---------------|----------------|------------------------|---------------|----------------|
| LambdaLR          | 0.6937        | 0.756          | CyclicLR - triangular  | 0.5427        | 0.66           |
| MultiplicativeLR  | 0.6953        | 0.753          | CyclicLR - triangular2 | 0.5430        | 0.677          |
| StepLR            | 0.5052        | 0.724          | CyclicLR - exp_range   | 0.5453        | 0.654          |
| MultiStepLR       | 0.6886        | 0.798          | OneCycleLR - cos       | 0.9456        | 0.756          |
| ExponentialLR     | 0.7778        | 0.763          | OneCycleLR - linear    | 1.1726        | 0.707          |
| CosineAnnealingLR | 0.5305        | 0.684          |                        |               |                |

Из результатов (таблица) видно, что относительно хорошее качество распознавания показали методы: LambdaLR (0.756), MultiplicativeLR (0.753), ExponentialLR (0.763), OneCycleLR – cos (0.756).

Лучшее качество показал метод уменьшения скорости обучения в определённые итерации цикла (MultiStepLR), 0.798.

### Литература

1. Pytorch.org. URL: <https://pytorch.org/> [Электронный ресурс]. (Дата обращения: 05.01.2021).
2. Руководство по планированию скорости обучения pytorch kaggle.com. URL: <https://www.kaggle.com/isbhargav/guide-to-pytorch-learning-rate-scheduling> [Электронный ресурс]. (Дата обращения: 15.01.2021).

**Гец Ольга Валерьевна**

Год рождения: 1996

Университет ИТМО,

Институт дизайна и урбанистики,

студент группы №42041,

направление подготовки: 09.04.03 – Прикладная информатика,

e-mail: zuzhuka21@gmail.com

**Иванов Сергей Владимирович**

Год рождения: 1980

Университет ИТМО,

Институт дизайна и урбанистики,

к.т.н., доцент,

e-mail: svivanov@itmo.ru

УДК 004.6

**ПРОЕКТИРОВАНИЕ ИТ-АРХИТЕКТУРЫ  
НА НЕФТЕДОБЫВАЮЩЕМ ПРЕДПРИЯТИИ**

**О.В. Гец**

**Научный руководитель – к.т.н., доцент С.В. Иванов**

**Аннотация**

В работе представлен анализ интеграции данных из нескольких систем-источников на нефтедобывающем предприятии, а также описаны дополнительные функции корпоративного хранилища данных, позволяющие снизить число ошибок, связанных с внутренними недостатками систем.

**Ключевые слова**

Обработка данных, корпоративное хранилище данных, иерархический справочник, мастер-система, нефтедобывающее предприятие.

Крупные нефтяные компании в настоящее время активно используют технологии, инструменты и библиотеки для обработки больших объемов данных, генерируемых их основной деятельностью.

Большие объемы данных в нефтегазовой сфере можно разделить на 2 класса: данные, получаемые из систем мониторинга и контроля технологических процессов, а также обработанные данные по управлению компанией и операциями.

Сложность организационной структуры, разрозненные производственные процессы на обширной территории и в самых разных сферах требуют решений от нефтегазовых компаний, позволяющих объединить все данные, необходимые для эффективного управления предприятием в едином информационном пространстве.

В результате проведенного исследования авторами работ был произведен анализ потоков данных из систем различного назначения в компании нефтегазовой отрасли. Перечень рассмотренных систем-источников и их основных характеристик, а также характеристик получаемых из них данных представлен в табл. 1.

Для интеграции данных из систем использовалось корпоративное хранилище данных на основе СУБД (системы управления базами данных) PostgreSQL. Модель хранилища была основана на типе MVCC (Multiversion Concurrency Control, многоверсионном управлении конкурентным доступом). Такая модель обеспечивает подключение к версии баз данных для каждого SQL-оператора на определенный момент времени, вне зависимости от текущего состояния данных. Такой способ

защищает операторы от несогласованности данных, возможной в случаях, если другие конкурирующие транзакции внесут изменения в те же строки данных, а также обеспечивает таким способом изоляцию транзакций для каждого сеанса баз данных.

Таблица 1

**Системы-источники данных**

| Система            | Описание системы  | Описание данных                          |
|--------------------|---|--|
| OIS Pipe           | Система сбора и анализа информации по промысловым трубопроводам                     | Технологические данные                   |
| OIS УСОИ           | Система контроля процесса добычи, переработки и сбыта углеводородов                 | Планово-экономические данные             |
| ИСУП               | Информационная система управления проектами   | Данные по проектам                       |
| Prognoz РИ ЗИП БРД | Система анализа прогнозных показателей разведки и добычи                            | Планово-экономические показатели         |
| 1С:ПБ              | Система автоматизации задач обеспечения охраны труда, производственной безопасности | Данные о производственной безопасности   |
| IBM Lotus Domino   | Система для организации среды совместной работы                                     | Локально-нормативные документы           |
| AZIMUT             | Система управления промышленной и экологической безопасностью, гражданской защитой  | Показатели производственной безопасности |
| MOSS ЕКП БРД       | Система контроля статусов крупных проектов  | Данные по проектам                       |

При выстраивании архитектурного взаимодействия было обнаружено несколько ключевых недостатков, влияющим на корректность интеграции данных и их дальнейшую визуализацию.

Одним из них являлось отсутствие первичного ключа Primary Key в некоторых системах, который позволил бы отличить актуальные записи от устаревших, а также производить инкрементальную загрузку только по тем записям, которые были изменены либо добавлены. Кроме того, при появлении новых записей в источнике по созданному в хранилище данных ключу актуальной может стать не самая последняя на момент транзакции, а любая новая, что вызывает недостаток в качестве данных, а впоследствии недоверие к ним, если подобная ошибка попадет на рабочие экраны лиц, принимающих решения. Для решения данной проблемы на этапе передачи данных из хранилища в использование системами Business Intelligence к каждому представлению были разработаны дополнительные служебные поля. Описание данных полей представлено в табл. 2.

Таблица 2

**Служебные поля хранилища данных**

| Наименование поля | Предназначение                     |
|-------------------|------------------------------------|
| Dwh_update_date   | Дата обновления                    |
| Dwh_change_id     | Идентификатор обновления           |
| Dwh_is_active     | Признак наличия записи в источнике |

Такое решение также позволило сохранить историю изменения значений одних и тех же показателей, что может являться предметом для анализа в каком-либо бизнес-процессе организации.

Кроме того, другим определяющим фактором для корректной интеграции являлось хранение данных на одинаковых уровнях иерархии для разных дочерних обществ или месторождений. В ходе анализа данных этот принцип был нарушен. Представление этого нарушения представлено на рисунке, где можно заметить, что в первом случае для дочернего общества существует 3 уровня детализации, а во втором случае – 4. Такие ситуации были выявлены как в разных системах, так и в одной и той же. Например, система OIS УСОИ хранила дополнительные предприятия одного дочернего общества на том же уровне, что и месторождения – для другого.

Как правило, в системах не выделяется уровень иерархии в дереве предприятий, поэтому при объединении систем вслепую может возникнуть расхождение. Его может обнаружить как пользователь, так и система верификации: как правило, расхождения начнутся с того уровня, где возникает отличие в иерархии, так что с каждым  $n$ -ым уровнем оно станет увеличиваться.

Для выявления этих расхождений функциональность хранилища данных была расширена алгоритмами сопоставления названий объектов (дочернее общество, месторождение, кустовая площадка, скважина и т.д.) из разных систем с названиями из выбранной мастер-системы для создания единого нормативного справочника в периметре исследуемых данных. По его итогу к каждому объекту был определен уровень иерархии, так что дальнейшая агрегация производилась с учетом найденных уровней (рисунок).

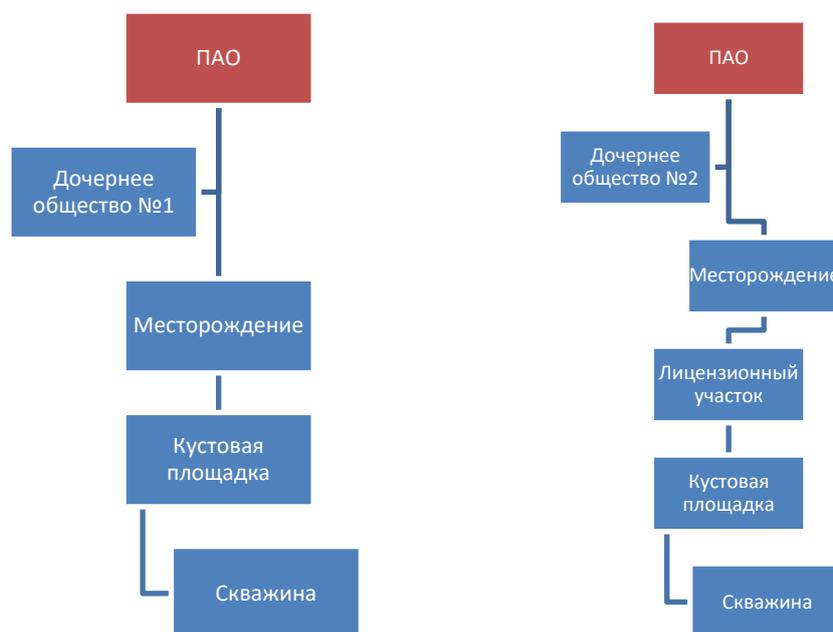


Рисунок. Сценарии иерархии данных

Данная проблематика и пути ее решения, в том числе, может относиться не только к нефтедобывающим предприятиям, но и к любой отрасли: например, разные ВУЗы имеют разную иерархию, и при необходимости показывать данные по нескольким университетам в данных возникнет много проблем. Например, в университете ИТМО есть деление на мегафакультеты, которые, в свою очередь, делятся на факультеты; однако у других ВУЗов, как правило, нет понятия «мегафакультетов».

Таким образом, выявленные проблемы систем-источников позволили расширить функционал корпоративного хранилища данных, что, в свою очередь, привело к увеличению качества данных и доверия лиц, принимающих решения, к интеграционным инструментам отображения информации. Визуализация данных из нескольких системных источников одновременно помогает компаниям расширять область знаний о всех ситуациях на производстве, в управлении и безопасности, что влечет за собой увеличение эффективности деятельности организации.



**Глухов Вадим Николаевич**  
Университет ИТМО,  
факультет информационных технологий и программирования,  
аспирант группы 7717,  
направление подготовки: 05.13.18 –Математическое  
моделирование, численные методы и комплексы программ,  
e-mail: vglukhov@itmo.ru



**Сметанников Иван Борисович**  
Университет ИТМО,  
факультет информационных технологий  
и программирования,  
к.т.н., доцент,  
e-mail: ismetannikov@itmo.ru

УДК 004.02

## **ОБЗОР ТЕКУЩЕГО СОСТОЯНИЯ МУЛЬТИВАРИАТИВНЫХ МЕТОДОВ ОТБОРА ПРИЗНАКОВ НА ОСНОВЕ ТЕОРИИ ИНФОРМАЦИИ**

**В.Н. Глухов**

**Научный руководитель – к.т.н., доцент И.Б. Сметанников**

### **Аннотация**

В работе рассмотрена эволюция методов отбора признаков, основанных на теории информации, начиная от наипростейших алгоритмов, не учитывающих избыточность признаков, до продвинутых алгоритмов, вводящих собственные определения релевантности и комплементарности между признаками, что позволяет учитывать не только простую взаимосвязь признака и целевой переменной по отдельности, но и находить сложные зависимости при взаимодействии признаков.

### **Ключевые слова**

Отбор признаков, фильтрующие алгоритмы, мультивариативные фильтры, теория информации.

Большинство задач машинного обучения, особенно в области биологии, медицины и генетики, имеют дело с большими объёмами данных, как в разрезе количества записей данных, так и в разрезе количества признаков, описывающих данные. Если большое количество записей является положительным фактором ввиду уменьшения смещения выборки от генеральной совокупности, то значительное количество признаков может оказать негативное влияние на решение задачи, что выражается в повышенном времени работы алгоритма и наличии незначимых признаков.

Проблему отбора признаков (feature selection) активно анализируют и развивают с начала 90-х годов XX века. Хотя существует большое количество разнообразных алгоритмов отбора признаков, группа алгоритмов, основанных на информационной теории, выглядит особенно привлекательной ввиду нескольких её преимуществ:

возможность измерения нелинейной взаимосвязи между признаками, а также работа как с категориальными, так и с интервальными признаками.

Прежде, чем переходить к самим методам, следует провести короткий экскурс в теорию информации и понятия, которыми оперирует проблема отбора признаков.

**Энтропия (entropy)** – мера хаоса данных. Энтропия также показывает вероятность полученного распределения данных случайной величины оказаться равномерным.

$$H(X) = -\sum_{x \in X} p(x) \log p(x).$$

**Условная энтропия (conditional entropy)** определяет энтропию в данных при условии новых знаний о данных.

$$H(X|Y) = -\sum_{x \in X} p(x|y) \log p(x|y).$$

**Взаимная информация (mutual information)** определяет прирост информации о переменной  $X$  после ввода новых данных в виде  $Y$ .

$$I(X; Y) = H(X) - H(X|Y).$$

Следующие определения имеют важное значение в теории отбора признаков.

**Релевантность (relevancy)** –  $I(X_k; Y)$ , свойство признака иметь объясняющую способность. Чем выше релевантность, тем ценнее признак, потому что он сильнее других уменьшает энтропию.

**Избыточность (redundancy)** –  $I(X_k; S)$ , свойство двух (и более) признаков нести общую информацию. Аналог из регрессионного анализа – мультиколлинеарность. Наличие мультиколлинеарности ухудшает регрессионную модель, увеличивая оценки.

**Взаимодополняемость (complementarity)** –  $I(X_k; S|Y)$ , эффект взаимодействия двух признаков, когда их наличие вместе дает больший эффект, чем сумма каждого из признаков отдельно

В определениях выше  $X_k$  – candidate feature, признак, рассматриваемый для включения в итоговый набор отобранных признаков;  $Y$  – целевая переменная;  $S$  – множество уже отобранных признаков.

Наиболее простой метод – Mutual Information Maximization – был предложен Lewis в 1992 году, который заключался в вычислении релевантности для каждого признака и отбора нескольких признаков с наибольшим показателем [1].

$$J_{MIM}(X_k) = I(X_k; Y).$$

Улучшением идеи занялся Battiti, который в 1994 году предложил следующую формулировку ранжирующего правила, названную Mutual Information Feature Selection [2]:

$$J_{MIFS}(X_k) = I(X_k; Y) - \beta \sum_{X_j \in S} I(X_k; X_j).$$

Он ввел штрафной коэффициент с гиперпараметром  $\beta$ , который уменьшает важность признака, если он сильно коррелирован с остальными. Гиперпараметр подбирается индивидуально от задачи. Автор в оригинальной статье предложил, что

единичный гиперпараметр будет достаточным в большинстве случаев, но теоретического объяснение этого представлено не было.

Yang и Moody в 1999 году предложили подход Joint Mutual Information, заключающийся в ранжировании признаков на основании следующей величины [3]:

$$J_{JMI}(X_k) = \sum_{X_j \in S} I(X_k X_j; Y).$$

Метрика метода представляет собой взаимную информацию, но не на одном признаке, а на совместном распределении уже отобранного признака и того, который потенциально собираются включить.

Informative Fragments предложен Vidal-Naquet и Ullman в 2003 году [4]. Он заключается в отыскании для каждого потенциального признака минимальной метрики, которая может быть записана как

$$J_{IF}(X_k) = \min_{X_j \in S} [I(X_k X_j; Y) - I(X_k; Y)].$$

Метрика имеет непростую интерпретацию, но при расписывании первого члена выражения в скобках может быть приведена к

$$J_{CMIM}(X_k) = \min_{X_j \in S} [I(X_k; Y|X_j)],$$

что является ничем иным, как выбором минимальной комплементарности для каждого потенциального признака [5]. Сам метод известен как Conditional Mutual Information Maximization.

Методом с наибольшим числом реализаций и дополнений является mRMR (Minimum Redundancy – Maximum Relevance), представленный в 2005 году, который представляет из себя измененную версию описанного выше алгоритма MIFS [6].

$$J_{mRMR}(X_k) = I(X_k; Y) - \frac{1}{|S|} \sum_{X_j \in S} I(X_k; X_j).$$

Таким образом, в данном подходе в штрафной коэффициент включается размер отобранного множества признаков, который побуждает отбирать большее число признаков, уменьшая тем самым штраф за избыточность.

Information Capping (ICAP), представленный Jakulin в 2005 году, развивает идеи предыдущих годов об особом представлении избыточности и выглядит следующим образом [7]:

$$J_{ICAP}(X_k) = I(X_k; Y) - \sum_{X_j \in S} \max [0, \{I(X_k; X_j) - I(X_k; X_j|Y)\}].$$

И, наконец, в 2015 году M.Bennasar et al развили идею JMI и представили Joint Mutual Information Maximization, где на каждой итерации отбирается наименьший коэффициент joint mutual information и выбираются несколько наилучших на основе полученного ряда [8]:

$$J_{JMIM}(X_k) = \min_{X_j \in S} [I(X_k X_j; Y)].$$

Таким образом, рассмотрены одни из наиболее важных представителей мультивариативных алгоритмов отбора признаков на основе теории информации. Многие из них послужили основой для более продвинутых алгоритмов, где

рассмотренные методы являются частью комплексных алгоритмов наряду с различными методами многомерной оптимизации. Знания того, как шло развитие алгоритмов, может стать подспорьем к открытию новых методов.

### Литература

1. Lewis D.D. Feature selection and feature extraction for text categorization. In Proceedings of the workshop on Speech and Natural Language, pages 212–217. Association for Computational Linguistics Morristown, NJ, USA, 1992.
2. Battiti R. Using mutual information for selecting features in supervised neural net learning. *IEEE Transactions on Neural Networks*, 5(4):537–550, 1994.
3. Yang H. and Moody J. Data visualization and feature selection: New algorithms for non-gaussian data. *Advances in Neural Information Processing Systems*, 12, 1999.
4. Vidal-Naquet M. and Ullman S. Object recognition with informative features and linear classification. *IEEE Conference on Computer Vision and Pattern Recognition*, 2003.
5. Fleuret F. Fast binary feature selection with conditional mutual information. *Journal of Machine Learning Research*, 5:1531–1555, 2004.
6. Peng H., Long F., and Ding C. Feature selection based on mutual information: Criteria of max-dependency, max-relevance, and min-redundancy. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 27(8):1226–1238, 2005.
7. Jakulin A. *Machine Learning Based on Attribute Interactions*. PhD thesis, University of Ljubljana, Slovenia, 2005.
8. Bennisar M., Hicks Y., Setchi R., Feature selection using Joint Mutual Information Maximisation, *Expert Systems with Applications*, Volume 42, Issue 22, 2015, Pages 8520-8532.



**Боченина Клавдия Олеговна**

Год рождения: 1988  
Университет ИТМО,  
факультет цифровых трансформаций,  
доцент, к.т.н.,  
e-mail: kbochenina@itmo.ru



**Гороховатский Леонид Юрьевич**

Год рождения: 1982  
Университет ИТМО,  
факультет цифровых трансформаций,  
доцент, к.псих.н.,  
e-mail: leongor\_us@yahoo.com

**Суриков Анатолий Георгиевич**

Университет ИТМО,  
национальный центр когнитивных разработок,  
к.т.н.,  
e-mail: anatoliy.surikov@itmo.ru

УДК 004; 159.9.072.5

**ТЕОРЕТИЧЕСКИЕ ПОДХОДЫ К ПРОБЛЕМЕ  
ИССЛЕДОВАНИЯ ЭМОЦИЙ В КИБЕРПРОСТРАНСТВЕ И ИХ  
ПЕРЕСЕЧЕНИЕ С КЛЮЧЕВЫМИ КОНЦЕПЦИЯМИ ЛИЧНОСТИ  
К.О. Боченина, Л.Ю. Гороховатский, А.Г. Суриков**

Работа выполнена в рамках темы НИР №620174 «Инкрементальное обучение моделей психоэмоциональной динамики пользователей и сообществ в цифровых социальных системах».

**Аннотация**

В работе рассмотрены феномен эмоций в контексте анализа деятельности актора киберпространства, подняты основные теоретические проблемы, возникающие в рамках исследований такого объекта на стыке психологии и компьютерных наук, сделаны некоторые теоретические обобщения, определены проблемы и перспективы исследований.

**Ключевые слова**

Эмоции, киберпространство, психология киберпространства.

Феномен эмоций сегодня представляется специалистам, проводящим свои изыскания на стыке психологии и компьютерных наук, одним из наиболее важных и перспективных с точки зрения будущего смыслового наполнения [8 и др.]. Это, вероятно, предопределено, в первую очередь, тем, что именно эмоции выступают наиболее точным «переходником», позволяющим перекинуть «мост» от исследования человека как носителя психики в аналоговой, условно объективной, реальности к

исследованию человека как актора киберпространства, а если говорить более точно – к исследованию актора киберпространства как носителя психики [3 и др.].

Эмоция – придерживаясь определения, данного этому понятию К.Э. Изардом, – есть «нечто, что переживается как чувство (feeling), которое мотивирует, организует и направляет восприятие, мышление и действие» [4]. Это нечто являет собой физиологически базирующееся явление. Основа любой эмоции (в этом сходятся без исключения все концептуализаторы рассматриваемого конструкта) – это физиологическая реакция человека на тот или другой раздражитель. Именно это обстоятельство обеспечивает условную объективность эмоциональных реакций, т.е. равнозначность эмоционального реагирования конкретного человека на конкретное рода раздражители как в аналоговой, так и в киберпространственной реакции.

Учитывая, что киберпространство даёт любому своему актору анонимность и относительную безответственность за свои действия (именно как принципиально не верифицируемому субъекту, скрывающемуся за маской своего аккаунта), исследователи активности киберпространственных сущностей нуждаются в инструментах отделения реальных личностей в сети от их цифровых клонов (роботизированных сущностей), либо от фальшивых сущностей (когда одна личность действует от имени другой, например, в случае закладывания чьим-либо реальным аккаунтом).

Другими словами: тем, кто хочет системно организовывать взаимодействие в киберпространстве на научной основе (будь то маркетинговая, политическая, социально-волонтерская, любая другая активность) необходимо чётко отделять реальные личности акторов киберпространства от так называемых «фейковых». Именно такую возможность даёт нам инструмент отслеживания эмоций.

Изучая эмоциональные реакции пользователей сети (в текстуальном, визуальном, аудиальном, семантически комбинированном проявлениях), мы имеем возможность чётко категоризировать акторов киберпространства по типам реальной личности, роботизированной сущности, фейковой личности.

Для осуществления такой категоризации мы должны чётко таксономизировать характерные эмоциональные проявления каждой из этой категории акторов.

Следующий важный этап познания – это отнесение типичных эмоциональных реакций акторов сетевой активности к одному из чётких и хорошо описанных в литературе психотипов. Установление соответствующей корреляции необходимо нам для того, чтобы прогнозировать поведение условно категоризованного как реального актора с его статистически характерными деятельностными особенностями и подходящим образом выстраивать взаимодействие с ним в интересах достижения наших целей.

Таким образом, перед современным исследователем стоит несколько задач: во-первых – это задача праксиметрического определения типов эмоционального реагирования акторов киберпространства; во-вторых – задача выработки алгоритма отделения реальных акторов сети от роботизированных и фейковых на основе эмоционального анализа; и, в-третьих, задача выявления взаимосвязи этих типов с некоторыми идеографически описанными типами личности (в целях прогнозирования поведения).

Важным теоретическим препятствием на пути решения этих задач является выбор адекватных концептуальных подходов к проблеме эмоций и их пересечение с наиболее подходящими ключевыми концепциями личности.

Проведя анализ большого количества концепций эмоций в современной психологии [1, 5, 7 и др.], мы пришли к выводу, что наиболее адекватная концепция эмоций, предлагаемая для нужд нашего исследования современной психологией – это концепция дифференциальных эмоций К.Э. Изарда [4].

Согласно К.Э. Изарду:

- базовые эмоции имеют отчетливые и специфические нервные субстраты;
- базовая эмоция проявляет себя при помощи выразительной и специфической конфигурации мышечных движений лица (мимики);
- базовая эмоция влечет за собой отчетливое и специфическое переживание, которое осознается человеком;
- базовые эмоции возникли в результате эволюционно-биологических процессов;
- базовая эмоция оказывает организующее и мотивирующее влияние на человека, служит его адаптации.

Приведенные особенности позволяют нам достаточно хорошо фиксировать эмоциональные проявления в киберпространстве (например, в социальных сетях).

Второй важный вопрос – это вопрос концепции личности, которая могла бы максимально четко быть связанной с предлагаемой концепцией эмоций.

В литературе [6] описано более 60 теорий личности. Вариант, показавшийся нам наиболее интересным, был найден в работе Е.П. Ильина [5].

Е.П. Ильин описывает концепцию, в рамках которой предлагается категоризировать личность по следующим параметрам: эмоциональность, активность и первичность.

Эмоциональность понимается как степень легкости возникновения эмоционального состояния в данной ситуации, а также его интенсивность.

Активность отражает потребность в действии, в осуществлении плана, замысла. Высокоактивные люди не выносят безделья, простоя. Поэтому они всегда находят себе занятие. Другие инертны, долго думают, прежде чем начать что-то делать.

Первичность характеризует другую сторону эмоциональности – быстроту исчезновения эмоции. Многие люди недолго остаются под впечатлением пережитого. Они непрестанно «разряжают» свои эмоции.

Исходя из сочетания этих свойств, были выделены восемь типов личности:

- нервный: Э+, А-, П;
- sentimentalный: Э+, А-, В;
- бурный, очень деятельный: Э+, А+, П;
- страстный: Э+, А+, В;
- сангвиник: Э-, А-, П;
- флегматик: Э-, А-, В;
- аморфный, или беспечный: Э-, А+, П;
- апатичный: Э-, А+, В.

Такой подход, как мы полагаем, наиболее адекватен обозначенной выше цели отнесения типичных эмоциональных реакций акторов сетевой активности к одному из чётких и хорошо описанных в литературе психотипов.

С учетом сказанного следует отметить, что праксиметрическое определение типов эмоционального реагирования акторов киберпространства должно осуществляться на основе концепции К.Э. Изарда, а задача выявления взаимосвязи этих типов с некоторыми идеографически описанными типами личности должна осуществляться с опорой на типы, описанные Е.П. Ильиным.

Важной проблемой остается проблема выработки алгоритма отделения реальных акторов сети от роботизированных и фейковых.

Размышляя над решением этой проблемы, мы констатировали, что имеем дело с серьезной методологической проблемой психологии киберпространства (как новой отрасли психологии). Суть проблемы укладывается в следующие тезисы:

- психология киберпространства – новая отрасль психологии;
- исследований в области психологии киберпространства сегодня – не более нескольких десятков [2 и др.];
- один из главных вопросов – это вопрос о предмете и объекте психологии киберпространства;
- основное внимание в существующих работах уделено влиянию компьютеров и интернета на жизнь человека;
- насколько совпадает, либо отличается активность личности в аналоговой и цифровой реальностях, т.е. можно ли говорить о том, что носитель психики в аналоговой реальности тождественен носителю психики в киберпространстве?

С учетом сказанного, мы определили для себя следующие пути научного поиска:

- построить методологическую схему типичного исследования психологического феномена в киберпространстве;
- апробировать эту схему;
- провести пилотные эксперименты на основе этой схемы.

### Литература

1. Вилюнас В.К. Психологические механизмы мотивации человека. М.: Изд-во Моск. ун-та. 1990.
2. Войскунский А.Е. Психология и Интернет. М.: Акрополь. 2010.
3. Гороховатский Л.Ю. Человек в цифровой экономике // Глобализация, цифровая трансформация, экономика, культура: новые специальности, новые горизонты: материалы XVI Международной научно-практической конференции. 24–26 октября 2018 года. СПб.: Издательство МБИ. 2018.
4. Изард К.Э. Психология эмоций. СПб.: Питер. 2007.
5. Ильин Е.П. Эмоции и чувства. СПб.: Питер. 2001.
6. Коржова Е.Ю. Психология личности: типология теоретических моделей. – Биробиджан, ИЦ ПГУ им. Шолом-Алейхема. 2015.
7. Экман П. Психология эмоций. СПб.: Питер. 2019.
8. Holyst J.A. & all. Cyberemotions. Collective Emotions in Cyberspace. Springer International Publishing, Switzerland. 2017.



**Гусев Алексей Евгеньевич**

Год рождения: 1995  
Университет ИТМО,  
факультет информационных технологий  
и программирования,  
аспирант,  
направление подготовки: 05.13.17 – Теоретические  
основы информатики,  
e-mail: roanlane@yandex.ru



**Волкова Марина Викторовна**

Год рождения: 1993  
Университет ИТМО,  
факультет информационных технологий  
и программирования,  
аспирант,  
направление подготовки: 05.13.15 – Вычислительные машины,  
комплексы и компьютерные сети,  
e-mail: volkova@speechpro.com



**Астапов Сергей Сергеевич**

Год рождения: 1988  
Университет ИТМО,  
факультет информационных технологий  
и программирования,  
ст.н.с., PhD,  
e-mail: sergeiastapov@gmail.com



**Кабаров Владимир Иосифович**

Год рождения: 1959  
Университет ИТМО,  
факультет информационных технологий  
и программирования,  
старший преподаватель,  
e-mail: vikabarov@itmo.ru

УДК 004.934.8'1

**ИССЛЕДОВАНИЕ МЕТОДОВ ГОЛОСОВОЙ  
БИОМЕТРИИ НА КОРОТКИХ ДЛИТЕЛЬНОСТЯХ  
А.Е. Гусев, М.В. Волкова, С.С. Астапов, В.И. Кабаров  
Научный руководитель – к.т.н., с.н.с. С.А. Новоселов**

Работа выполнена в рамках темы НИР №718574 «Методы, модели и технологии искусственного интеллекта в биоинформатике, социальных медиа, киберфизических, биометрических и речевых системах».

### **Аннотация**

В данной работе рассматривается задача текстонезависимой верификации диктора на коротких произнесениях в задаче IVR. Анализируется применение глубоких нейронных сетей на базе архитектур ECAPA-TDNN и ResNet для построения голосовых моделей в различных акустических условиях.

### **Ключевые слова**

Верификация и идентификация диктора, голосовая биометрия, глубокие нейронные сети.

Стремление современных компаний автоматизировать рутинные процессы и повысить скорость обработки клиентских запросов стимулировало развитие голосовых роботизированных систем [1]. Один из примеров обработки входящих телефонных звонков – так называемые интерактивные голосовые меню (IVR, Interactive Voice Response). IVR-системы позволяют перенаправлять звонки в автоматическом режиме, предоставлять необходимую информацию по запросу пользователя и тем самым существенно снизить нагрузку на операторов контакт-центра. Изначально IVR-системы ожидали ответа пользователя в виде тонального набора, однако развитие методов машинного обучения позволило внедрять в IVR-системы технологии распознавания речи и голосовую биометрию.

К технологии голосовой биометрии в таком случае предъявляется требование верифицировать пользователя на фразах короткой длительности, например, «фамилия - имя - отчество». При этом не должно быть ухудшения качества при верификации на длинных фразах. В данной работе мы сравнили различные виды нейросетевых систем верификации диктора и протестировали их в сценариях «короткий тест» - «длинный эталон» и «длинный тест» - «длинный эталон» на данных в телефонном канале.

Глубокие нейронные сети последние несколько лет стали новым стандартом в системах верификации диктора по голосу. Обучаясь задаче классификации на большом количестве известных дикторов, такие нейронные сети учатся экстрагировать информацию о дикторе из входных акустических признаков и аккумулировать её в одном из последних слоёв, выход которого называют эмбедингом (embedding). Задача верификации сводится к сравнению близости эталонного и тестового эмбедингов, например, по косинусной метрике или с использованием дополнительных техник сравнения.

Наиболее распространённые на сегодняшний архитектуры экстракторов используют блоки с одномерными (TDNN) или двумерными (ResNet) свёртками. Для возможности обработки сигналов произвольной длительности после свёрточных слоёв используется слой статистического пулинга (statistics pooling layer), возвращающий вектор фиксированной размерности.

В работе [2] была предложена архитектура ECAPA-TDNN, развивающая идею x-vector TDNN архитектуры с добавлением SE-Res2Block. Схема архитектуры представлена на Рисунке 1a. Входной вектор признаков подаётся на Conv1D (TDNN) слой, затем проходит через три блока SE-Res2Block, выходы которых конкатенируются перед подачей на следующий Conv1D слой.

Каждый SE-Res2Block устроен согласно рисунку 1b. В нем используется два Conv1D слоя с параметром delay >1, что позволяет захватывать разный временной контекст. Центральный Res2Net модуль представляет собой иерархическую систему фильтров для выделения признаков разных масштабов (рисунок 1c). SE (Squeeze-Excitation) блока реализует механизм поканального внимания, перевзвешивая выход Conv1D блока.

Конкатенированные выходы трёх SE-Res2Block подаются на Conv1D слой, далее применяется Attentive statistics pooling, который находит взвешенное среднее и

взвешенное стандартное отклонение последовательности высокоразмерных карт признаков, полученных с предыдущего блока сети. Следующий полносвязный слой возвращает векторное представление диктора, последний линейный слой имеет размерность по числу классов и используется для обучения сети (рисунок).

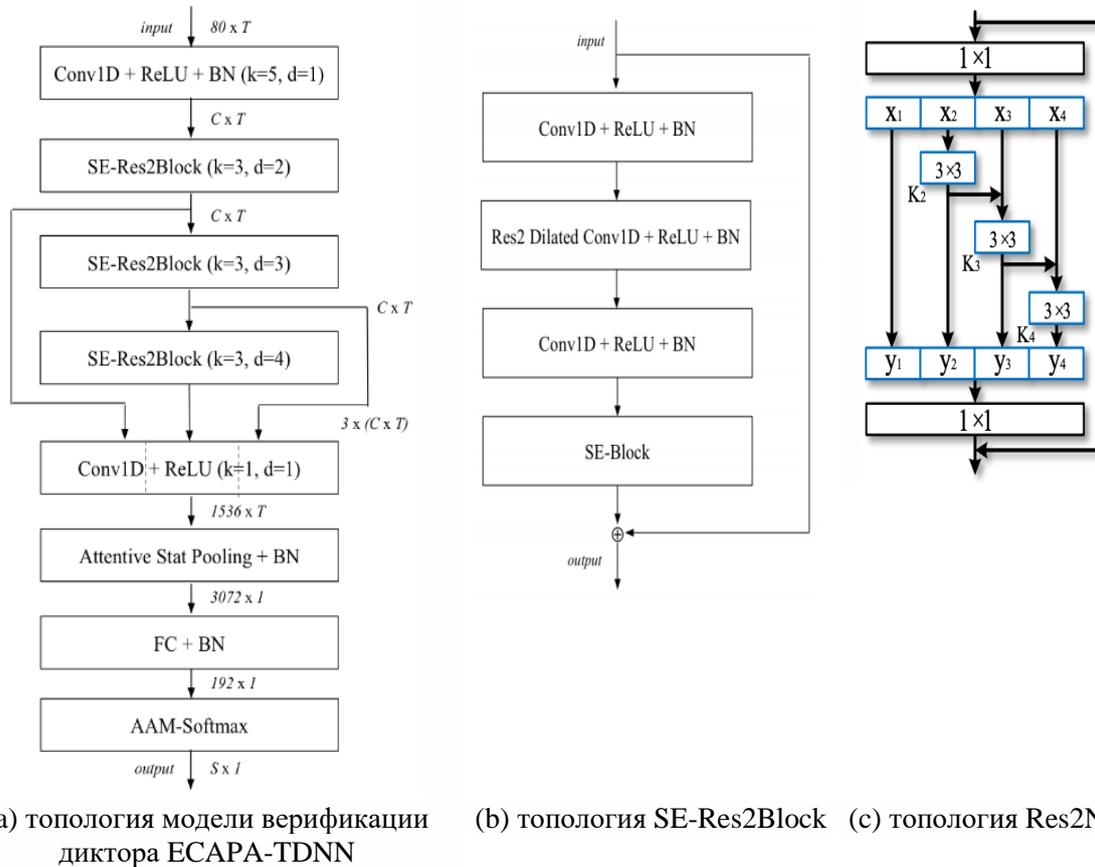


Рис. 1. Архитектура ECAPA-TDNN

Для обучения использовались датасеты из конкурсов по распознаванию диктора (NIST-SRE, VoxCeleb1-2), а также собственные базы STC-Telecom и обучающее множество IVR-STC. Для тестирования на коротких длительностях использовались фразы «фамилия-имя-отчество» тестовой базы IVR-STC, записанной в разнообразных шумовых условиях. Для тестирования на файлах по 10-60 секунд речи использовалась база конкурса NIST SRE19-CTS-eval-set. Эталонная модель в обоих случаях строилась на 60 и более секундах речи.

Для подготовки обучающих данных применялась аугментация шумами базы MUSAN (music, babble, noise) и реверберацией. В качестве входных признаков рассматривались 23-MFCC и 64-MelFilerBanks с применением нейросетевого детектора речи [3] и локальной нормализацией. Обучающие примеры были разбиты по 6 секунд, при дообучении на длинных фрагментах (в Таблице обозначены как 'tune') – по 12-18 секунд. В качестве базовых архитектур использовались TDNN (x-vector), ResNet различной глубины (34, 101) и ECAPA-TDNN. Использовались функции потерь: AM-Softmax с параметрами  $s=30$ ,  $m=0.2-0.4$  и CurricularFace [4] с параметрами  $s=30$ ,  $m=0.2-0.4$ . Дополнительно применялся метод нормализации оценок сравнений записей (в Таблице обозначено, как 's-norm'). При тесте на длинных длительностях также произведено сравнение с SOTA архитектурой, обученной в рамках [5] (в таблице обозначен  $j$  как 'ResNet34-JHU'). Результаты представлены в таблице.

**Результаты экспериментов**

|                              | IVR-STC |                         | SRE19-CTS-eval-set |                              |
|------------------------------|---------|-------------------------|--------------------|------------------------------|
|                              | EER(%)  | minDCF <sub>0.001</sub> | EER(%)             | minDCF <sub>0.01-0.005</sub> |
| X-vector                     | 12.05   | 0.975                   | ---                | ---                          |
| ECAPA-TDNN-tune              | 8.44    | 0.869                   | 3.33               | 0.360                        |
| ResNet34-AM                  | 5.82    | 0.834                   | 5.69               | 0.530                        |
| ResNet34-AM-tune             | 6.01    | 0.858                   | 4.38               | 0.450                        |
| ResNet101-AM-tune            | 6.09    | 0.816                   | 3.29               | 0.332                        |
| ResNet101-CF-tune            | 5.68    | 0.804                   | 2.78               | 0.300                        |
| ResNet101-CF-s-<br>norm-tune | 6.83    | 0.822                   | 2.52               | 0.270                        |
| ResNet34-JHU                 | ---     | ---                     | 2.68               | 0.248                        |

Из результатов, представленных в Таблице, можно сделать следующие выводы:

- для верификации на коротких файлах (IVR-CTS) модели на conv2d блоках (ResNet) превосходят по качеству модели с conv1d (TDNN);
- использование дообучения на длинных сегментах и нормализации скоров повышает качество работы системы на длинных файлах и слабо влияет на качество на коротких файлах;
- более глубокие нейронные сети показывают лучшее качество;
- CurricularFace позволяет обучить более качественную модель, чем AM-Sortmax.

В работе было проведено сравнение актуальных на сегодняшний день нейросетевых архитектур для голосовой биометрии (TDNN, ResNet, ECAPA-TDNN) и их модификаций для задачи верификации диктора на коротких длительностях в телефонном канале. Было показано, что архитектуры на основе ResNet показывают лучшее качество в сценарии IVR, при этом дообучение на длинных сегментах и использование нормализации скоров позволяет сохранить качество работы на длинных файлах. Это даёт возможность использовать унифицированную модель голосовой биометрии как для задач IVR, так и для разговоров с оператором произвольной длительности.

**Литература**

1. <https://voximplant.ru/blog/avtomatizacia-kontakt-centrov-v-kompaniah-elektronnoj-torgovli> (дата обращения: 22.02.2021).
2. Brecht Desplanques, Jenthe Thienpondt, Kris Demuynck, "ECAPA-TDNN: Emphasized Channel Attention, Propagation and Aggregation in TDNN Based Speaker Verification», arXiv:2005.07143v3 [eess.AS] 10 Aug 2020.
3. Lavrentyeva G., Volkova M., Avdeeva A., Novoselov S., Gorlanov A., Andzukaev T., Ivanov A., Kozlov A. "Blind speech signal quality estimation for speaker verification systems" , Proceedings of the Annual Conference of the International Speech Communication Association, INTERSPEECH – 2020. pp. 1535-1539.

4. Huang Y., Wang Y., Tai Y., Liu X., Shen P., Li S., Li J., and Huang F. "Curricularface: adaptive curriculum learning loss for deep face recognition". In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, pages 5901–5910, 2020.
5. Villalba J., Garcia-Romero D. et al "Advances in Speaker Recognition for Telephone and Audio-Visual Data: the JHU-MIT Submission for NIST SRE19", Odyssey 2020.



**Двойникова Анастасия Александровна**

Год рождения: 1996  
ФГБУН «Санкт-Петербургский Федеральный исследовательский центр Российской академии наук» (СПб ФИЦ РАН),  
программист,  
Университет ИТМО,  
факультет информационных технологий и программирования,  
студент группы № М42211,  
направление подготовки: 09.04.02 – Информационные системы и технологии,  
e-mail: dvoynikova.a@iias.spb.su



**Мамонтов Данила Юрьевич**

Год рождения: 1994  
Университет Ульма, Ульм, Германия,  
аспирант,  
Университет ИТМО,  
факультет информационных технологий и программирования,  
аспирант,  
e-mail: danila.mamontov@uni-ulm.de



**Карпов Алексей Анатольевич**

Год рождения: 1978  
ФГБУН «Санкт-Петербургский Федеральный исследовательский центр Российской академии наук» (СПб ФИЦ РАН),  
д.т.н., доц., главный науч. сотрудник,  
e-mail: karpov@iias.spb.su

УДК 004.934.2

**АВТОМАТИЧЕСКОЕ ОПРЕДЕЛЕНИЕ ЭМОЦИОНАЛЬНОГО СОСТОЯНИЯ УЧАСТНИКОВ ПРЕДМЕТНЫХ РАЗГОВОРОВ ПО ТРАНСКРИПЦИЯМ РЕЧИ**

**А.А. Двойникова, Д.Ю. Мамонтов, А.А. Карпов**  
**Научный руководитель – д.т.н., доцент А.А. Карпов**

Работа выполнена при поддержке РФФ (проект №18-11-00145).

**Аннотация**

В работе проводятся экспериментальные исследования по определению уровня эмоциональных проявлений в текстовых транскрипциях базы данных K-EmoCop. Рассматривается влияние сбалансирования классов при обучении классификаторов на точность определения эмоций. В статье устанавливается базовый стандарт результатов по классификации уровня эмоций дикторов в текстовых транскрипциях.

**Ключевые слова**

Распознавание эмоций, анализ речевых транскрипций, дисбаланс классов, сбалансирование классов, определение уровня эмоций.

## **Введение**

Выражение эмоциональной экспрессии через речевые высказывания происходит с помощью двух модальностей: акустической и лингвистической. Зачастую многие исследователи анализируют только акустическую модальность, исключая лингвистическую составляющую [1]. Лингвистическая модальность передает смысловое содержание речевых высказываний, через которое активно выражаются чувства, эмоции, мнения дикторов. Поэтому анализ лингвистической информации разговорной речи является основополагающим для определения эмоций в натуралистических разговорах.

Проблема дисбаланса классов является актуальной в машинном обучении в различных задачах. Как показывают предыдущие исследования [2, 3] решение проблемы дисбаланса классов позволяет улучшить точность классификации.

В работе используется мультимодальная база K-EmoCon [4]. На сегодняшний день не известны работы с исследованиями этой базы данных, поэтому результаты, представленные в этой статье, можно считать baseline для анализа эмоций по текстовым транскрипциям.

## **База данных**

В экспериментальных исследованиях использовалась база данных K-EmoCon [4]. Это мультимодальная англоязычная база данных, содержащая аудио, видео и физиологические сигналы дикторов, которые общались в парах и обсуждали статью на тему йеменских беженцев. Перед записью базы данных участникам предлагалось прочитать статью и выписать аргументы, отражающие их отношение к ситуации, поэтому речь не была импровизированной. Каждый диалог длился 10 минут, 2 минуты говорил один диктор, потом следующие 2 минуты – другой, затем снова менялись. Всего в базе данных содержится 16 аудиозаписей диалогов между 32 участниками, из которых 20 мужчин и 12 женщин возрастом от 19 до 36 лет. База данных K-EmoCon записывалась в Корее, соответственно участниками базы данных являлись корейцы, к которым были выдвинуты следующие требования: они должны грамотно говорить по-английски (IELTS уровень разговорного языка не ниже 7), но не являться носителями языка, также иметь не менее трех лет опыта проживания в англоязычной стране.

Аннотирование диалогов происходило каждые 5 секунд по следующим категориям: полярность и интенсивность эмоций (5 уровней), 5 основных эмоций – счастье, радость, грусть, нервозность, злость (4 уровня) и 11 аффективных категорий (2 класса). Аудиозаписи аннотировались с 3 сторон: себя (участник размечал проявление своих экспрессий), партнера и внешней (5 сторонних аннотаторов).

## **Предобработка базы данных**

Как описывалось выше, при записи базы данных было установлено правило, что каждый собеседник должен говорить 2 минуты, потом передавать слово другому. Но в ходе записи участники могли выйти за заданный тайминг. Поэтому для выделения речи каждого диктора из диалогов проводилась диаризация дикторов вручную. Аудиозаписи базы данных K-EmoCon нарезались на кусочки, каждый из которых содержал речь только одного диктора, всего из 16 аудиозаписей получилось 153 кусочка.

Так как авторы базы данных не предоставили транскрипции речевых выражений собеседников, необходимо было извлечь текстовую информацию диалогов из аудиозаписей. Извлечение транскрипций происходило с помощью автоматической системы распознавания речи (англ. Automatic Speech Recognition, ASR) Silero [5]. Данная ASR позволяет распознавать речь на английском, немецком и испанском языках с помощью предобученных моделей [6]. Из некоторых аудио кусочков не удалось извлечь текстовые транскрипции, т.к. они содержали невнятную речь или были

сильно зашумлены. Также из базы данных удалялись транскрипции длиной меньше 2 слов, потому что в таких коротких высказываниях нет проявления эмоциональных экспрессий. Объем базы данных K-EmoCon составил 143 текстовых транскрипций. Так как разметка базы данных происходила по аудиозаписям каждые 5 секунд, а речевые фразы одного диктора длились дольше одного временного интервала разметки, то для проставления эмоциональных меток в транскрипциях применялся следующий алгоритм. Для каждой транскрипции записывались все соответствующие уровням эмоции метки аннотаторов в интервале от начала до конца речевого высказывания для каждой отдельной эмоции, подсчитывалась частота встречаемости каждой метки, и текстовой транскрипции присваивалась та метка, которая была наиболее встречаемая. Если для одной транскрипции получалось одинаковое количество двух разных меток, то транскрипции присваивалась та метка, которая соответствовала наивысшему уровню эмоции. Помимо распознавания базовых эмоций, также определялось проявление любой эмоциональности, для этого все транскрипции сгруппировывались в 2 класса – нейтральный и эмоциональный. К нейтральному классу относились те транскрипции, у которых полярность и интенсивность эмоции равнялась 3 (среднему) уровню, к эмоциональному классу относились все транскрипции, отличные от нейтрального класса. В данной работе исследовалась аннотирование только одной стороны: самого себя каждого участника. На рис. 1 представлено распределение транскрипций для каждой категории отдельно, в соответствии с ее уровнем.

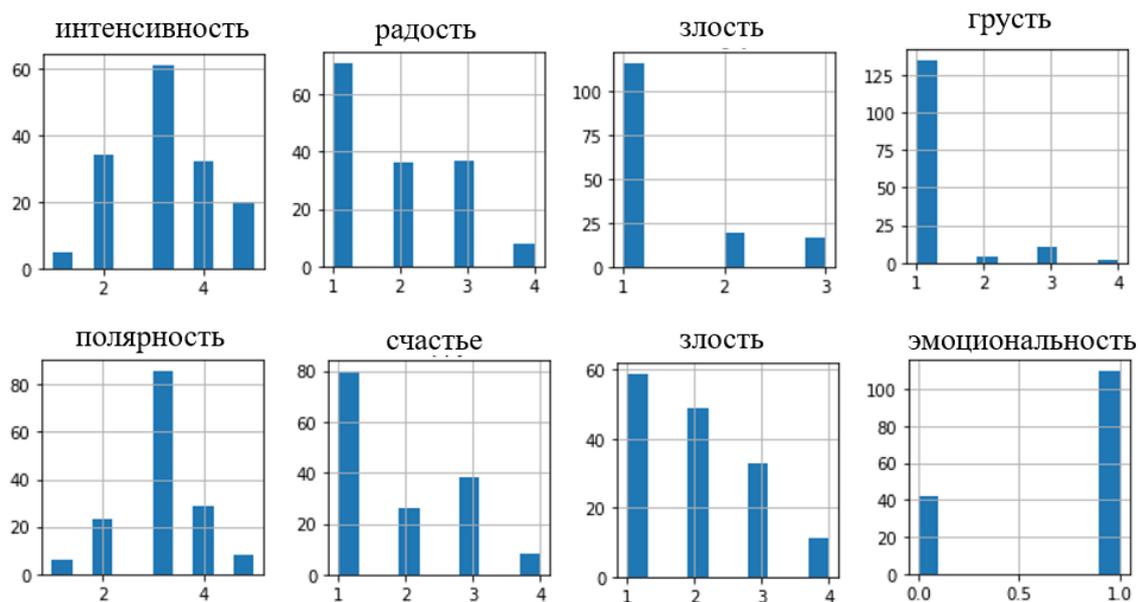


Рис. 1. Распределение транскрипций для каждой категории

Следующим этапом предобработки данных являлась обработка текстовых транскрипций. Полученные текстовые транскрипции содержали в себе синтаксические и орфографические ошибки, это связано с тем, что система автоматического распознавания речи работает не всегда корректно. Поэтому из транскрипций вручную удалялись повторяющиеся слова, добавлялись и удалялись пробелы там, где это необходимо. Далее проводилась токенизация слов, удаление пунктуации и стоп-слов (например, «а», «an», «the», «do» и т.п.), удаление неправильно распознанных слов, список таких слов составлялся вручную и включал такие слова как «dge», «est», «gd», «mle», «umi», «ye» и т.п. Заключительным этапом предобработки транскрипций являлась лемматизация, она изменяет форму слова в соответствии с его начальной формой, тем самым происходит нормализация всех слов.

### Экспериментальные исследования

Текстовые транскрипции базы данных K-EmoCon имеют разметку по 5 базовым эмоциям – счастье, радость, злость, нервозность и грусть. Для каждой эмоции строились отдельные классификаторы, определяющие уровень эмоции, также отдельно строился бинарный классификатор для определения наличия эмоциональности. В экспериментальных исследованиях, представленных в данной работе, использовались следующие классификаторы: логистическая регрессия (англ. Logistic Regression, LR), наивный Байес (англ. Naive Bayes, NB), метод опорных векторов (англ. Support Vector Machine, SVM) и случайный лес (англ. Random Forest, RF). Перед тем как построить классификаторы, необходимо векторизовать текст. Векторизация текста происходила с помощью следующих методов: мешок слов (англ. Bag-of-Words, BoW), Word2Vec, FastText, BERT. Для последних трех методов использовались предобученные модели. Информация о длине векторов и предобученных моделях для каждого метода представлена в табл. 1.

Таблица 1

Информация о методах векторизации текстовых транскрипций

| Метод    | Длина вектора | Предобученная модель               |  |
|----------|---------------|------------------------------------|--|
|          |               | Используемая модель                | Корпус данных  |
| BoW      | 1161          | -                                  | -  |
| Word2Vec | 300           | GoogleNews-vectors-negative300 [7] | Корпус новостей Google. Объем около 100 млрд. слов               |
| FastText | 300           | wiki-news-300d-1M.vec [8]          | Корпус обучен на англоязычной Википедии. Объем около 1 млн. слов |
| BERT     | 768           | bert-base-uncased [9]              | Корпус обучен на отзывах с сайта Amazon. Объем около 8 млн. слов |

Все текстовые транскрипции базы данных K-EmoCon делились на обучающий и тестовый набор в процентном соотношении 70/30 для каждой эмоции отдельно. Обучение классификаторов происходило с помощью 3-х блочной кросс-валидации, для каждого классификатора подбирались наиболее эффективные параметры. На рис. 2 представлена диаграмма точности классификации по количественному показателю невзвешенной средней полноты (англ. unweighted average recall, UAR), определения уровня эмоции. Для каждой эмоции использовалась комбинация описанных выше методов векторизации и классификации, которая показывала наибольшее значение UAR.

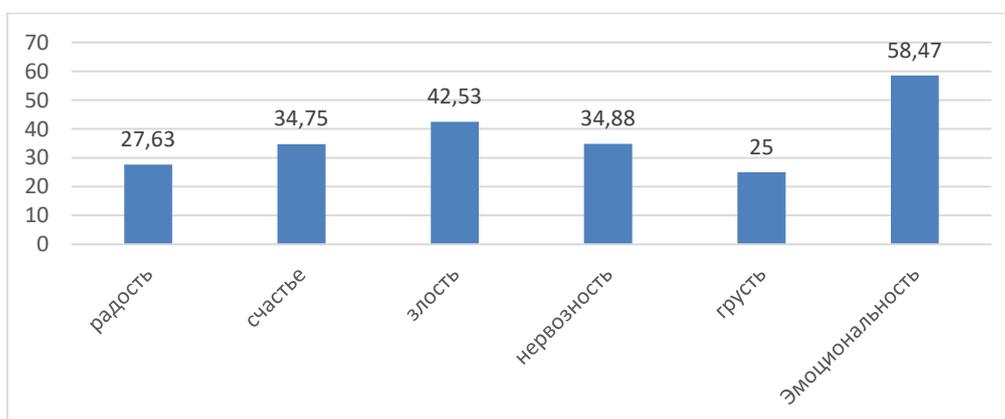


Рис. 2. Качество распознавания уровня каждой эмоции по показателю UAR, %

Как видно из рис. 1, данные для каждой эмоции являются несбалансированными. Чтобы решить проблему дисбаланса классов, предлагается применить взвешивание классов. В библиотеке `scikit-learn` [10] языка программирования Python для классификаторов можно установить функцию, которая позволяет назначать определенный вес каждому классу, в соответствии с формулой (1).

$$coefficient = \frac{n\_samples}{n\_classes * np.bincount(y)}, \quad (1)$$

где *coefficient* – коэффициент изменения весов, *n\_samples* – количество примеров в одном классе, *n\_classes* – количество классов, *np.bincount* – функция, которая подсчитывает количество вхождений каждого значения в массив *y*, *y* – массив меток.

Применение данной функции позволяет произвести автоматическую настройку весов обратно пропорционально частотам классов, т.е. наименьшему классу присваивается больший вес, и наоборот.

В табл. 2 представлены результаты классификации уровней для каждой эмоции в соответствии с применением балансирования классов и без нее. Также представлена информация о наиболее эффективной комбинации методов векторизации и классификации для каждой эмоции, результаты работы которых выражены метрикой UAR.

Таблица 2

**Результаты классификации уровня каждой эмоции по показателю UAR, %**

| Эмоция          | Без сбалансирования |              | Со сбалансированием |               | Разница UAR, % |
|-----------------|---------------------|--------------|---------------------|---------------|----------------|
|                 | UAR                 | Методы       | UAR                 | Методы        |                |
| Радость         | 27,63               | BERT, RF     | 31,57               | BoW, RF       | 3,94           |
| Счастье         | 34,75               | FastText, RF | 33,05               | BoW, LR       | -1,7           |
| Злость          | 42,53               | Bert, SVM    | 54,24               | Bert, SVM     | 9,71           |
| Нервозность     | 34,88               | Bert, LR     | 45,48               | Bert, SVM     | 10,6           |
| Грусть          | 25,00               | Bert, SVM    | 59,62               | Bert, SVM     | 34,62          |
| Эмоциональность | 58,47               | Fasttext, NB | 66,53               | Fasttext, SVM | 8,06           |

Как видно из результатов, представленных в таблице 2, сбалансирование классов позволяет повысить точность распознавания уровня эмоций для каждой эмоции, кроме счастья. Также можно заметить, что комбинация метода векторизации BERT и метода опорных векторов в качестве классификатора позволяет достичь наилучшего результата среди других комбинаций.

### Вывод

В работе анализировалось определение уровня эмоций в транскрипциях базы данных K-EmoCon, полученных с помощью автоматической системы распознавания речи. Также рассматривалось влияние балансировки классов на точность распознавания, как показывают экспериментальные исследования, балансировка классов при распознавании уровня эмоций в текстовых транскрипциях позволяет повысить значение показателя UAR с 3,94 % до 34,62 % отдельно для каждой эмоции, кроме эмоции счастья, для которой сбалансирование классов уменьшило значение UAR на 1,7 %. Использование метода векторизации BERT и метода опорных векторов во многих экспериментах позволяет получить наивысший результат по классификации уровней эмоций. В данной работе удалось достичь наибольшей точности определения уровня эмоций для радости UAR = 31,57 %, счастья – 34,75 %, злости – 54,24 %, грусти – 59,62 %, эмоциональности – 66,53 %.

нервозности – 45,48 %, грусти – 59,62 % и определения эмоциональности – 66,53 %. Данные результаты можно считать базовым стандартом (baseline) для последующих экспериментов с текстовыми транскрипциями базы данных K-EmoCon.

### Литература

1. Dvoynikova A., Verkholyak O., Karpov A. Emotion recognition and sentiment analysis of extemporaneous speech transcriptions in Russian // Lecture notes in computer science. SPECOM 2020. 2020. Vol. 12335 LNAI. pp. 136–144.
2. Ширяев А.И., Нехаев И.Н. Двухэтапный метод классификации писем для решения задачи технической поддержки клиентов // Гибридные и синергетические интеллектуальные системы. 2018. С. 317-322.
3. Баймаханбетов М.А. и др. ADASYN как метод решения проблемы несбалансированности данных при сентимент-анализе текстовой информации // Современное педагогическое образование. 2019. №. 12.
4. Park C.Y. et al. K-EmoCon, a multimodal sensor dataset for continuous emotion recognition in naturalistic conversations // Scientific Data. 2020. Т. 7. №. 1. С. 1-16.
5. Silero Speech [Электронный ресурс]. Режим доступа: <https://www.silero.ai/> (дата обращения: 25.02.2021).
6. Silero Models [Электронный ресурс]. Режим доступа: <https://github.com/snakers4/silero-models> (дата обращения: 25.02.2021).
7. Google Code word2vec [Электронный ресурс]. Режим доступа: <https://code.google.com/archive/p/word2vec/> (дата обращения: 25.02.2021).
8. FastText English word vectors [Электронный ресурс]. Режим доступа: <https://fasttext.cc/docs/en/english-vectors.html> (дата обращения: 25.02.2021).
9. Pytorch pretrained bert 0.6.2 [Электронный ресурс]. Режим доступа: <https://pytorch.org/project/pytorch-pretrained-bert/> (дата обращения: 25.02.2021).
10. Scikit-learn Machine Learning in Python [Электронный ресурс]. Режим доступа: <https://www.sklearn.org/> (дата обращения: 25.02.2021).



**Дудина Дарина Сергеевна**  
Университет ИТМО,  
факультет инфокоммуникационных технологий,  
студент группы №К41111с,  
направление подготовки: 11.04.02 – Инфокоммуникационные  
технологии и системы связи,  
e-mail: dudinadarina@gmail.com



**Басов Олег Олегович**  
Университет ИТМО,  
факультет инфокоммуникационных технологий,  
д.т.н., доцент,  
e-mail: oobasov@itmo.ru

УДК 004.021

## **АНАЛИЗ СПОСОБОВ И СРЕДСТВ ИЗВЛЕЧЕНИЯ И ИНТЕРПРЕТАЦИИ ПАРАМЕТРОВ КЛАВИАТУРНОГО ПОЧЕРКА ПОЛЬЗОВАТЕЛЯ**

**Д.С. Дудина**

**Научный руководитель – д.т.н., доцент О.О. Басов**

Работа выполнена в рамках темы НИР №620185 «Разработка интеллектуальной многопользовательской адаптивной системы освещения для снижения конфликтного поведения и стрессовых ситуаций в образовательной среде».

### **Аннотация**

В работе рассматриваются программные и аппаратные средства извлечения параметров клавиатурного почерка, а также математические подходы, которые используются для задач аутентификации, идентификации пользователя по клавиатурному почерку. Сделан вывод о возможности использования рассмотренных средств и методов в дальнейших разработках, направленных на определение психоэмоционального состояния пользователя.

### **Ключевые слова**

Клавиатурный почерк, средства извлечения параметров клавиатурного почерка, распознавание эмоций, идентификация, аутентификация.

Все люди имеют отличительные особенности, например отпечатки пальцев, подобная особенность есть и в цифровом пространстве. Временные параметры работы пользователя на клавиатуре являются стабильным признаком, который проявляется на подсознательном уровне. Когда человек набирает текст на клавиатуре, у него задействуется порядка пятидесяти различных мышц одновременно, и каждый человек их использует по-разному и с разной силой [1]. Сделать вывод об обучаемости человека можно на основе того, как быстро он ищет нужные клавиши. Время

удержания клавиши длится несколько микросекунд, но также является уникальным действием.

Клавиатурный почерк представляет собой поведенческую биометрическую характеристику, описывающую такие параметры, как:

1. Динамика ввода – это время между нажатиями клавиш и временем их удержания.

2. Скорость ввода – это количество напечатанных символов, разделенное на время печати.

3. Использование клавиш для печати заглавных букв.

4. Частота возникновения ошибок при вводе [2].

Один из способов обнаружения эмоций пользователя – это анализ шаблонов клавиатурного почерка пользователя и типа текста (слова, предложения), вводимого им. Такой комбинированный анализ демонстрирует значительное количество эмоциональных состояний, обнаруженных при вводе данных пользователем.

Несколько алгоритмов машинного обучения были использованы для анализа атрибутов времени нажатия клавиш и текстового шаблона. Нажатие клавиши – это самый дешевый и доступный способ взаимодействия с компьютерами. Были рассмотрены семь эмоциональных классов для классификации эмоциональных состояний. Для анализа текстовых шаблонов использовалась модель векторного пространства с мерой сходства Жаккара для классификации ввода произвольного текста. Данный комбинированный подход показал точность определения эмоций более 80% [3].

### **Средства извлечения параметров клавиатурного почерка**

Все средства извлечения параметров клавиатурного почерка могут быть разделены на следующие группы.

Акустические средства – это маленькие устройства с встроенным микрофоном и системой хранения информации. Данные устройства позволяют восстановить текст, напечатанный на телефоне, и осуществляют его передачу. Свободный оборот этих средств запрещён.

Аппаратные средства – это устройства, которые подключаются в разъём PS/2 или USB клавиатуры и представляют собой небольшие устройства со встроенной памятью или со сменной картой памяти. Также у них могут присутствовать модули Bluetooth или Wi-Fi, что позволяет сохранять полученные данные и передавать их в файловые хранилища.

Преимущества:

- a. легко установить и подключить устройство;
- b. не требует дополнительного питания;
- c. легко проверить работоспособность.

Недостатки:

a. пользователь может заметить, что за ним осуществляется мониторинг и может быть более осторожным в общении;

b. устройство или карта памяти может быть извлечена пользователем, чтобы этого избежать нужно опломбировать разъёмы и проверять поступление информации;

c. данные средства не обладают функцией перехвата нажатий клавиш мыши и перемещения её курсора и выполнение скриншотов.

### **Программные средства**

Классификация программных средств извлечения параметров клавиатурного почерка представлена на рисунке.

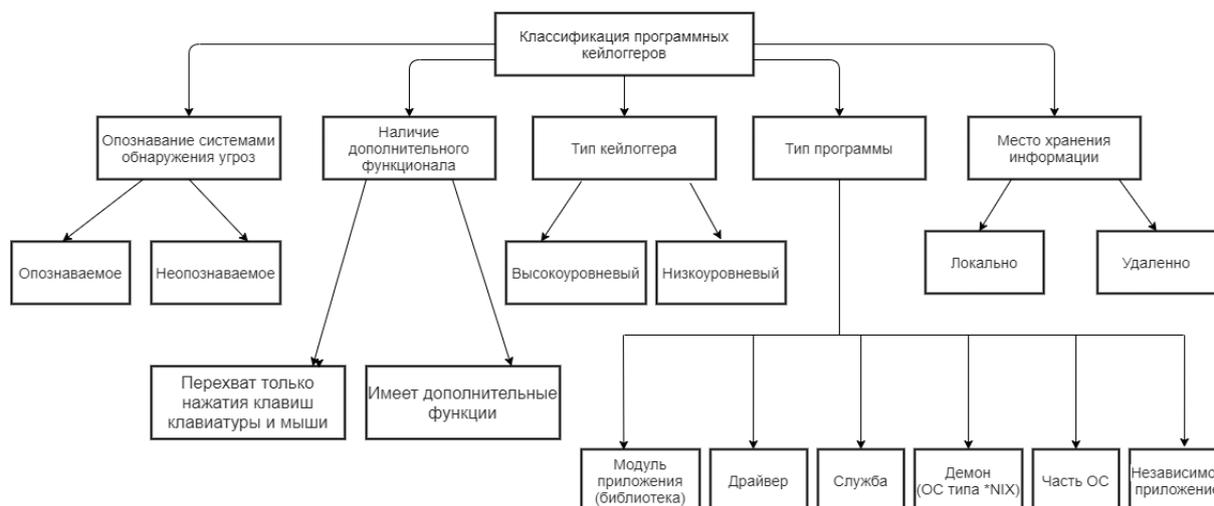


Рисунок. Классификация программных средств извлечения параметров клавиатурного почерка

Рассмотрим подробнее виды программных средств:

- 1) По типу опознавания антивирусами и детекторами уязвимостей средства могут быть опознаваемыми или неопознаваемыми;
- 2) По наличию дополнительных функций программные средства могут повторять аппаратные или имеют функции журнала работы пользователя в виде скриншотов и фиксировать приложения, в которые вводился текст;
- 3) Программное средство может быть высокоуровневым (перехватывать только текст, который был введен в приложение) или низкоуровневым, то есть перехватывает и нажатия управляющих и модифицирующих клавиш;
- 4) По типу программы программные средства могут быть приложением, драйвером, службой, «демоном», частью операционной системы, независимым самостоятельным приложением;
- 5) По типу хранения перехваченной информации можно разделить на локальное или удалённое хранение. Целесообразнее пользоваться двумя методами: если сервер доступен, то полученная информация отправляется периодически, если нет — сохраняется локально.

### Обзор алгоритмов распознавания клавиатурного почерка

Характерным отличием таких алгоритмов распознавания клавиатурного почерка является наличие режима обучения, то есть пользователь должен пройти аутентификацию и ввести какой-нибудь текст. Далее происходит считывание динамических характеристик пользователя и их последующее сохранение. Таким образом, после обучения в системе происходит накопление данных о времени удержания клавиш и паузах между нажатиями для всех известных пользователей.

Разные авторы предлагают различные способы хранения данных, например в виде вектора, который содержит набор длительностей событий клавиатуры (удержание клавиши, наложение нажатий, пауза между нажатиями) [4]. Также данные могут быть представлены в виде массива, элементами которого могут быть тип клавиши, вид события (нажатие или отпускание клавиши) и время события [5].

Можно выделить несколько групп алгоритмов распознавания клавиатурного почерка:

1. Алгоритмы, которые анализируют почерк при вводе пароля и обладают высоким быстродействием, потому что пользователю нужно ввести только пароль, но если пароль будет коротким, то точность будет низкой.

2. Алгоритмы, которые анализируют почерк после ввода дополнительного текста, являются более точными, по сравнению с предыдущими алгоритмами. Недостатком является то, что ввод дополнительного текста требует больше времени, что может привести к появлению негативных эмоций у пользователя.

3. Алгоритмы, осуществляющие постоянно скрытый мониторинг параметров клавиатурного почерка, обеспечивают высокую точность, но потребляют больше ресурсов.

Для задач аутентификации, идентификации и распознавания клавиатурного почерка можно выделить следующие математические подходы:

- 1) статистические;
- 2) вероятностно-статистические;
- 3) на основе распознавания образов и нечеткой логики;
- 4) нейронные сети.

Статистический подход заключается в том, что пользователю необходимо зарегистрироваться в системе, потом происходит сбор статистики в виде вычисляемых параметров особенностей динамики работы на клавиатуре. Далее статистические данные усредняются и получается биометрический эталон, который характеризует нового пользователя. При аутентификации пользователь вводит ключевую фразу для того, чтобы она сравнивалась с полученным ранее эталоном. Сравнение происходит путем вычисления введенной ключевой фразы с помощью какого-либо критерия меры близости и биометрического эталона, в основном этот подход используют при аутентификации по парольной фразе.

Вероятностно-статистический подход основан на теории вероятности и математической статистики. Он заключается в том, что время нажатия и отпускания клавиш при печати рассматриваются как вероятностные события. Предполагается, что значения времени удержания и пауз между удержаниями клавиш распределены по закону, который можно приближенно считать нормальным законом распределения.

Математический подход на основе распознавания образов и нечеткой логики для решения задач распознавания клавиатурного почерка довольно распространен. Большинство таких алгоритмов базируются на решающем правиле Байеса и его модификациях.

Нейронные сети слабо распространены для решения поставленных задач по сравнению с предыдущими подходами. Так как часто время обучения нейронных сетей заранее неизвестно и достаточно велико [6].

Исходя из изученных материалов, оценка психофизиологического и эмоционального состояний пользователя с помощью анализа клавиатурного почерка и движения мыши может быть информативной, а существующие алгоритмы идентификации, аутентификации пользователя по клавиатурному почерку могут быть применены в целях определения состояния пользователя, но требуется проведение работ по поиску новых способов выделения информативных признаков, характерных паттернов и эффективных методов их распознавания. Это в будущем позволит разработать программное средство, которое определяло бы психофизиологическое состояние пользователя и в будущем создать систему, способную определять начальные признаки изменения психоэмоционального состояния.

## Литература

1. Иванов А.И., Биометрическая идентификация личности по динамике подсознательных движений. / Иванов А.И. Пенза: Изд.-во Пенз. Гос. Ун-та. 2000. 188 с.

2. Ponen J. Keystroke Dynamics [Электронный ресурс]. Lappeenranta University of Technology. 2008. Режим доступа: <http://researchweb.iiit.ac.in/~vandana/PAPERS/KS/Ponen.pdf> (дата обращения: 08.01.2020).
3. Nahin A.F.M.N.H., Alam J.M., Mahmud H., and Hasan K. Identifying emotion by keystroke dynamics and text pattern analysis. Behaviour & Information Technology, 33(9):987–996, 2014.
4. Гистограммный метод распознавания клавиатурного почерка. Брюхомицкий Ю.А. Известия Южного федерального университета. Технические науки. 2010. Т.112. №11. стр. 55-58.
5. Три алгоритма управления доступом к КСИИ на основе распознавания клавиатурного почерка оператора. Сидоркина И.Г., Савинов А.Н. Вестник Чувашского университета. 2013. №3. стр. 293-301.
6. Ишмурзина И.У. Исследование и разработка методологии идентификации обучающегося методом клавиатурного почерка в системах дистанционного обучения: диплом. работа. Санкт-Петербургский Политехнический Университет Петра Великого, Санкт-Петербург. 2017.



**Евсеева Елена Сергеевна**

Год рождения: 1997  
Университет ИТМО,  
факультет информационных технологий и программирования,  
студент группы №М42211,  
направление подготовки: 09.04.02 – Информационные  
системы и технологии,  
e-mail: evseevalenoshka@yandex.ru



**Астапов Сергей Сергеевич**

Год рождения: 1988  
Университет ИТМО,  
факультет информационных технологий и программирования,  
с.н.с., PhD,  
e-mail: sergeiastapov@gmail.com

УДК 004.93

**ОЦЕНКА ПРИМЕНИМОСТИ МОДЕЛИ  
COUNTNET ДЛЯ ПОДСЧЕТА ЧИСЛА ДИКТОРОВ**

**Е.С. Евсеева**

**Научный руководитель – с.н.с., PhD С.С. Астапов**

Работа выполнена в рамках темы НИР №620172 «Определение структуры полилога с применением лексических и нелексических признаков речи нескольких дикторов».

**Аннотация**

В данной работе приведен обзор существующих решений и подходов по определению мгновенного числа одновременно говорящих дикторов на записях с дальнего микрофона, а также протестировано готовое решение CountNet на одном из открытых наборов данных для различных задач классификации.

**Ключевые слова**

Обработка речевых сигналов, подсчет числа дикторов, рекуррентные нейронные сети, CountNet, AMI Corpus.

Определение мгновенного числа одновременно говорящих дикторов – это задача определения числа дикторов на аудиозаписи в каждый момент времени. Обычно для этого аудиозапись нарезается на фрагменты определенной длины, и для каждого из них определяется число активных дикторов. Данное знание важно при решении задач разделения дикторов и выделения целевого диктора [1, 2]. Также число дикторов полезно знать для приложений, анализирующих речь, повышая их устойчивость в ситуациях, когда активно несколько дикторов. К таким приложениям относятся распознавание речи, мониторинг толпы, обнаружение конфликтов и др. Для них известное число одновременно говорящих дикторов оказывается очень важным для эффективности последующих процессов.

Цель данной работы заключается в оценке применимости модели CountNet для задачи подсчета числа дикторов. В рамках данной работы приведен обзор существующих решений и подходов по определению мгновенного числа одновременно говорящих дикторов на записях с дальнего микрофона, а также рассмотрено готовое

решение CountNet и протестировано на одном из открытых наборов данных для различных задач классификации.

Рассмотрим часто используемые архитектуры и методы для решения задачи определения числа дикторов. В качестве архитектуры решений часто используются сверточные нейронные сети (CNN) или рекуррентные нейронные сети (RNN) [3]. CNN и RNN успешно используются в системах распознавания речи, обнаружения перекрывающейся речи, а также распознавания дикторов. Для предобработки речевых сигналов и выделения из них признаков используются такие методы, как: огибающая сигнала, гистограмма речевого сигнала, частотный спектр, спектрограмма, мел-частотные кепстральные коэффициенты (MFCC).

Рассмотрим используемые наборы данных. В качестве наборов для обучения часто используются открытые базы данных, такие как AMI Corpus [4] и LibriSpeech [5]. AMI Corpus – это набор данных, состоящий из 100 часов записей совещаний групп, преимущественно из четырех человек, на английском языке. Корпус AMI подходит в качестве данных для тестирования в поставленной задаче, так как он состоит из реальных записей с дальнего микрофона. Однако он хуже проявляет себя в обучении моделей, так как в большинстве времени на записях активно по одному диктору. Частым решением такой проблемы является создание синтетических данных, в основном на основе корпуса LibriSpeech. Данный корпус создан на основе аудиокниг и содержит несколько наборов данных суммарно на 1000 часов записей зачитанной английской речи. Объединение записей различных дикторов дает возможность создавать огромное количество входных выборок для обучения моделей.

Одним из существующих готовых решений для данной задачи является модель CountNet [6]. Авторы предлагают современную аудиомодель глубокой нейронной сети, основанную на двунаправленной архитектуре сети долгой краткосрочной памяти (Bi-LSTM) для оценки количества дикторов. Для ее обучения авторами был создан синтетический набор данных LibriCount10 [7] на основе наборов данных LibriSpeech clean-360 и акустических сцен TUT. Набор содержит 5720 сгенерированных уникальных смесей дикторов длительностью пять секунд для тестового набора с максимальным количеством одновременно говорящих дикторов равным 10. Для представления речевого сигнала используется кратковременное преобразование Фурье, вычисленное с использованием окон Ханна. На вход модели подаются короткие звуковые фрагменты длительностью 5 секунд для предсказания количества дикторов путем классификации фрагментов на 11 классов: от 0 до 10 дикторов.

Для тестирования модели CountNet возьмем AMI корпус. В табл. 1 представлено количество дикторов, встречающееся в данном корпусе, и сколько таких примеров встречается в корпусе на отрезках записей по 5 секунд. По таблице видно, что в AMI корпусе присутствует 5 классов по количеству дикторов, и самые многочисленные из них – где присутствует 1 и 2 диктора.

Таблица 1

#### Кол-во примеров в корпусе AMI

| Кол-во дикторов       | 0    | 1     | 2     | 3    | 4   |
|-----------------------|------|-------|-------|------|-----|
| Кол-во отрезков аудио | 3972 | 33697 | 27682 | 6147 | 453 |

Произведем тестирование как на оригинальных записях корпуса, так и на дополнительно усиленных на 20 дБ, так как записи в основном имеют низкую энергию. Результат тестирования по метрике F1-score для обычных данных составил 0,240, для усиленных – 0,422. Большой прирост результата на усиленных данных свидетельствует

о том, что, для модели сложно распознавать записи с дальнего микрофона, которые имеют малую мощность сигнала.

С целью получить прирост результата протестируем модель еще для 3-х задач классификации: от 0 до 2+ дикторов, то есть включающая в себя классы 0, 1, 2+ диктора, от 0 до 3+ и от 0 до 4+ дикторов. Результаты для данных задач по метрике F1-score приведены в табл. 2.

Таблица 2

### Результаты тестирования задач классификации на корпусе АМІ

| Задачи классификации | F1-score       |                  |
|----------------------|----------------|------------------|
|                      | Обычные данные | Усиленные данные |
| 2+                   | 0,283          | 0,544            |
| 3+                   | 0,242          | 0,430            |
| 4+                   | 0,240          | 0,422            |

Результат закономерно растет с уменьшением количества классов. Лучший достигается на задаче от 0 до 2+ дикторов на усиленных данных и равен 0,544 по метрике F1-score. Для данного результата на рисунке представлены матрицы спутанности с метриками recall и precision. Precision можно интерпретировать как долю объектов, названных классификатором положительными и при этом действительно являющимися положительными, а recall показывает, какую долю объектов положительного класса из всех объектов положительного класса нашла модель. На рисунке видно, что модель склонна недооценивать количество дикторов. Наиболее сильно это проявляется для количества дикторов равного 2+, помечая такие примеры чаще как активность 1 диктора (рисунок).

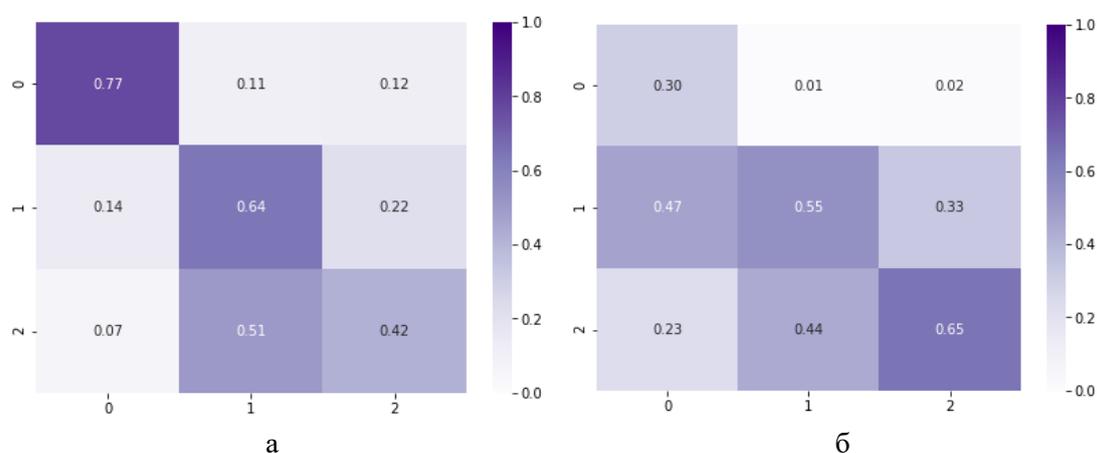


Рисунок. Матрицы спутанности с метриками: recall (а); precision (б)

В данной работе приведен обзор решений задачи определения мгновенного числа дикторов, рассмотрено готовое решение CountNet и протестировано для различных задач классификации. Модель CountNet можно использовать для задачи классификации от 0 до 2+ дикторов, но при этом реальные данные требуют предобработки в виде усиления сигнала. Также большой недостаток заключается в том, что она использует длинные фрагменты аудио по 5 секунд, что усложняет ее использование в реальных ситуациях. Дальнейшее исследование состоит в рассмотрении и применении моделей на основе других архитектур для обработки фрагментов аудио меньшей длительности.

### Литература

1. Astapov S., Lavrentyev A., Shuranov E., «Far Field Speech Enhancement at Low SNR in Presence of Nonstationary Noise Based on Spectral Masking and MVDR Beamforming». In: Karpov A., Jokisch O., Potapova R. (eds) // Speech and Computer. SPECOM 2018. Lecture Notes in Computer Science, vol 11096. 2018. Springer, pp. 21-31.
2. Astapov S., Popov D., Kabarov V., «Directional Clustering with Polyharmonic Phase Estimation for Enhanced Speaker Localization». In: Karpov A., Potapova R. (eds) // Speech and Computer. SPECOM 2020. Lecture Notes in Computer Science, vol 12335. 2020. Springer, pp. 45-56.
3. Andrei V., Cucu H., Burileanu C., «Overlapped Speech Detection and Competing Speaker Counting—Humans Versus Deep Learning» // IEEE Journal of Selected Topics in Signal Processing. 2019. vol. 13, no. 4, pp. 850-862.
4. Carletta J., «Announcing the AMI Meeting Corpus» // The ELRA Newsletter 11(1). 2006. P. 3-5.
5. Panayotov V., Chen G., Povey D., and Khudanpur S. Librispeech: «An ASR corpus based on public domain audio books» // In Proc. IEEE (ICASSP). 2015. P. 5206–5210.
6. Stöter F-R. et al. «Classification vs. Regression in Supervised Learning for Single Channel Speaker Count Estimation» // IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP). 2018. P. 436-440.
7. Stöter F-R. et al. «LibriCount, a dataset for speaker count estimation» // Zenodo. 2018.

**Емельянова Валентина Игоревна**

Университет ИТМО,  
Мегафакультет трансляционных информационных технологий,  
Институт дизайна и урбанистики,  
студент группы № С42011,  
направление подготовки: 27.04.07 – Наукоемкие технологии и экономика инноваций,  
e-mail: emelianova.valentina97@gmail.com

**Якимук Илья Витальевич**

Университет ИТМО,  
Мегафакультет трансляционных информационных технологий,  
Институт дизайна и урбанистики  
аспирант группы № 7738,  
направление подготовки: 05.13.10 – Управление в социальных  
и экономических системах,  
e-mail: ilyayakimuk@gmail.com

**Митягин Сергей Александрович**

Университет ИТМО,  
Мегафакультет трансляционных информационных технологий,  
Институт дизайна и урбанистики,  
к.т.н.,  
e-mail: mityagin@itmo.ru

УДК 004.023

**МЕТОД ОЦЕНКИ ПОТЕНЦИАЛА  
НАБЕРЕЖНЫХ КАК ДОСУГОВЫХ ПРОСТРАНСТВ  
В.И. Емельянова, И.В. Якимук  
Научный руководитель – к.т.н., доцент С.А. Митягин**

Работа выполнена в рамках темы НИР №620181 «Средства определения тематики и значимости объектов городской среды на основе открытых данных социальных сетей».

**Аннотация**

В работе рассмотрен метод оценки набережных, позволяющий определить потенциал использования прибрежных территорий как досуговых пространств. Проведено экспериментальное исследование, показывающее корректность и эффективность работы метода оценки набережных, находящихся в черте города. Оценка проводится по трем основным критериям: общие качественные характеристики, характеристики качества досуговой среды и видовые характеристики территории. В результате исследования получена оценка потенциала проведения на набережной того или иного вида досуга.

**Ключевые слова**

Городская набережная, оценка потенциала, метод, досуговое пространство, инфраструктура, социальные группы, открытые данные.

Горожане, зачастую, не знают, чего ожидать от определенного пространства, его функционального наполнения и качества в целом, какой формат досуга они смогут получить, попадая на определенную территорию, чем и обосновывается актуальность данного исследования. Подобрать место отдыха в соответствии со своими запросами о планируемом досуге, получить максимальное удовлетворение от посещения того или

инного пространства является обоснованным запросом современного городского жителя. Так как прибрежные территории занимают значительную часть Санкт-Петербурга, но также значительно обделены вниманием к существующему потенциалу, жителю сложнее узнать достоверную информацию о качестве набережных как досуговых пространств [1]. Таким образом, в данной работе проводится экспериментальное исследование метода оценки потенциала набережных как досуговых пространств.

Методика определения пригодности набережной к определенному формату подразумевает:

1. Выбор исследуемого объекта.
2. Сбор данных о выбранном объекте. Сбор данных осуществляется из различных источников. Первостепенными являются сервисы открытых данных. К ним относятся общедоступные сервисы, такие как Open Street Map и Wikimapia, сервисы с ограниченным доступом – РГИС. Помимо существующих общедоступных данных, необходимо провести полевые исследования для сбора недостающей, более точечной информации и социологические исследования в виде социологических опросов населения.
  3. Анализ полученных данных по выделенным критериям:
    - а) общие качественные характеристики:
      - транспортно-пешеходный каркас;
      - вид береговой кромки;
      - вид контактной зоны;
      - открытость объекта;
    - б) характеристики качества досуговой среды (инфраструктурное наполнение):
      - качество покрытий;
      - зеленый каркас территории;
      - наличие инфраструктуры;
      - освещение территории;
    - с) видовые характеристики:
      - общая просматриваемость территории: объекты культурного наследия, значимые доминанты;
        - природные особенности. Солнце;
        - природные особенности. Луна.
  4. Присвоение оценки объекту исследования по шкале пригодности к определенному формату досуга (рисунок).

В рамках данной работы проводится экспериментальное исследование наиболее значимой набережной города Санкт-Петербурга - набережной реки Фонтанка. Для присвоения общей итоговой оценки необходимо провести анализ территории в соответствии с описанным методом.

Общие качественные характеристики определяются на основе натурных исследований в соответствии с разработанной типологией от MLA+, описанной в исследовании «Нераскрытый Петербург. Вода» [2]. Береговая линия набережной считается основной отличительной чертой пространства набережной от любого другого общественного городского пространства. Внешний вид, качество и возможность использовать воду и берег в полной мере позволяют привлекать в пространство как можно больше посетителей, поэтому оценка состояния береговой линии набережной является неотъемлемой для составления индивидуальной комплексной оценки [3].

Качество досуговой среды – важнейший комплекс показателей, который является ключевым в определении принадлежности территории общественного пространства к виду досуговой деятельности. Этот критерий позволяет оценить степень

интеграции набережной в систему городской зелёной инфраструктуры (состоящей из разнообразных общественных пространств и природных ландшафтов), а также конструктивную основу экоустойчивого современного прибрежного объекта [4, 5]. Приспособленность физического состояния покрытий к возможности передвижения различных групп населения. Своеобразие и индивидуальность инфраструктурного наполнения, а также уровень безопасности общественного пространства.

Из результатов видно, что набережная реки Фонтанки рекомендована для туристической прогулки ( $D=3,3$ ), прогулки на велосипеде ( $D=3,4$ ). К возможному формату досуга можно отнести прогулку ( $D=1,7$ ). Нерекондуемыми форматами досуга являются прогулка с детьми ( $D=0,9$ ) и пляжный отдых ( $D=0,4$ ).

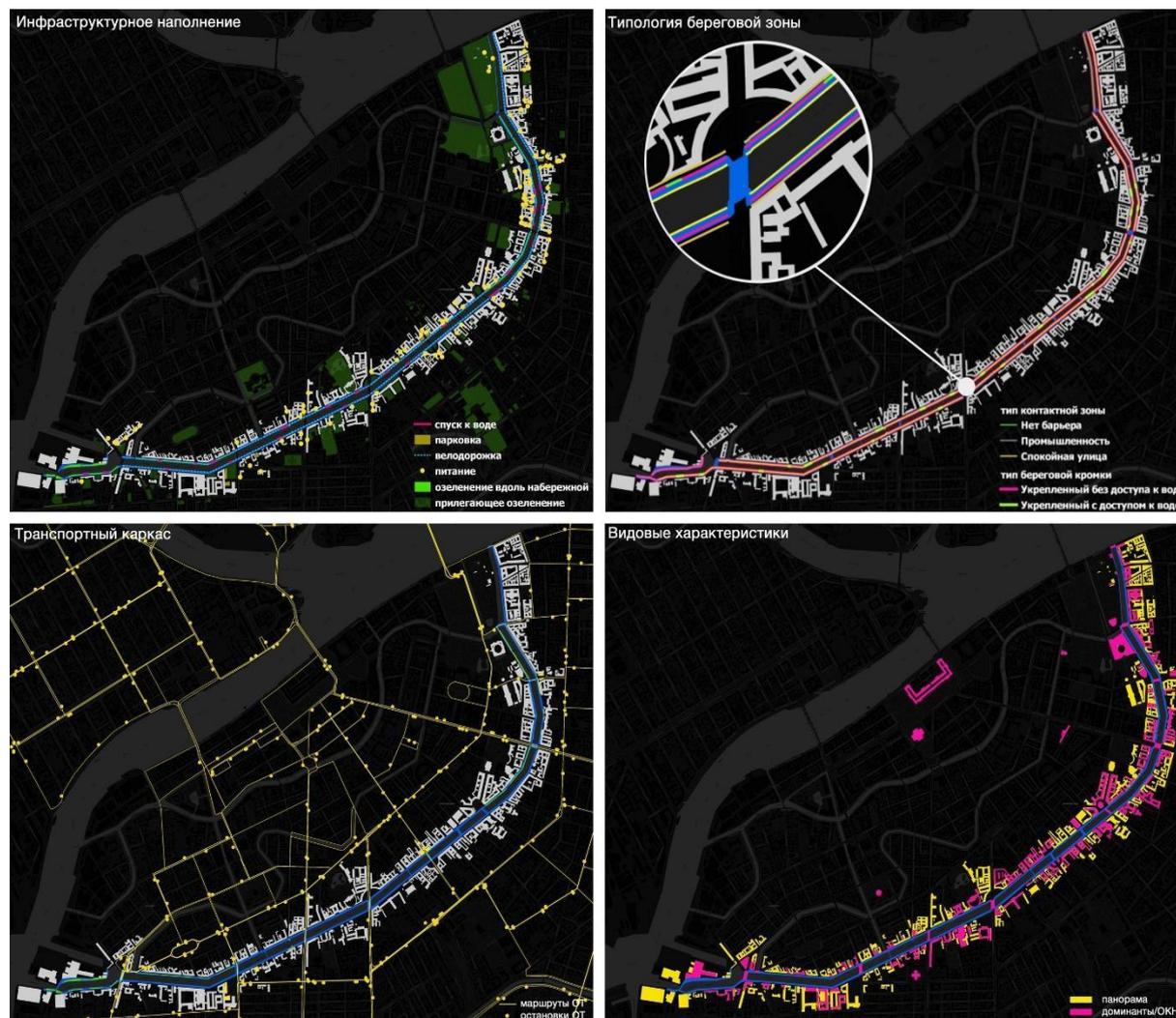


Рисунок. Этапы методики определения пригодности набережной

Отличительной чертой представленного метода является наличие оценки видовых качеств территории. Сюда включены как характеристики прилегающей панорамы застройки, высотные доминанты, так и природный аспект (таблица).

Совокупность оценок по описанным выше показателям интегрирована в итоговый расчет оценки потенциала набережной, производящийся по формуле:

$$D = \sum d * M. \quad (1)$$

**Результаты эксперимента**

| Критерий                      | Пляжный<br>отдых | Туристическая<br>прогулка | Велосипед | Прогулка<br>с детьми | Прогулка |
|-------------------------------|------------------|---------------------------|-----------|----------------------|----------|
| Транспортно-пешеходный каркас | -0,2             | 0,1                       | 0,4       | -0,2                 | 0,075    |
| Вид береговой кромки          | -0,458           | 0,106                     | 0,0312    | 0,21                 | 0,459    |
| Вид контактной зоны           | 0,276            | 0,08                      | 0,0675    | 0,2                  | 0,0735   |
| Качество покрытий             | 0,183            | 0,158                     | 0,338     | 0,278                | 0,175    |
| Зеленый каркас                | 0,5              | 0,1                       | 0,1       | 0,1                  | 0,125    |
| Инфраструктура                | 0,1              | 0,225                     | 1,7525    | 0                    | 0,125    |
| Освещение                     | 0                | 0,2                       | 0,5       | 0,25                 | 0,25     |
| Общая просматриваемость       | 0                | 2                         | 0,18      | 0,06                 | 0,3      |
| Солнце                        | 0                | 0,1                       | 0,02      | 0,0125               | 0,05     |
| Луна                          | 0                | 0,225                     | 0,045     | 0,0225               | 0,1125   |
| Оценка (D)                    | 0,401            | 3,294                     | 3,4342    | 0,933                | 1,745    |

В ходе работы было проведено экспериментальное исследование метода оценки потенциала набережных как досуговых пространств. Провелась работа с различными источниками данных, а также проведен ряд натуральных и социологических исследований, на основе которых собраны данные, необходимые для анализа выбранной набережной. Проведен анализ данных по критериям, описанным в методике оценки. Данные по каждому критерию визуализированы и сведены в общую таблицу. Проведен расчет оценки в соответствии с критериями и выведенными коэффициентами. Полученный результат оценки набережной реки Фонтанки не противоречит действительности, поэтому представленный метод считается успешным для дальнейшего испытания на других набережных. Разработанный метод позволит повысить качество жизни горожан и туристов и позволит привлечь больше посетителей на набережные.

В дальнейшей работе планируется применение метода на ряде набережных города Санкт-Петербурга, с целью подтверждения эффективности метода, уточнения критериев и исправления выявленных недостатков.

**Литература**

1. Эра пост-цифровых технологий уже наступила. [Электронный ресурс] // РОСКОНГРЕСС: [сайт]. URL: <https://roscongress.org/materials/era-post-tsifrovyykh-tekhnologiy-uzhe-nastupila-a-vy-gotovy-k-budushchemu/> (дата обращения: 17.01.2021).
2. MLA+. Нераскрытый Петербург. Вода 2019. Режим доступа: [https://www.mlaplus.com/wp-content/uploads/2019/10/Undiscovered-St.Petersburg\\_water\\_MLA.pdf](https://www.mlaplus.com/wp-content/uploads/2019/10/Undiscovered-St.Petersburg_water_MLA.pdf) (дата обращения: 09.01.2021).
3. Городская среда [Электронный ресурс] // Минстрой России: [сайт]. [2020]. Режим доступа: <https://www.minstroyrf.ru/trades/gorodskaya-sreda/> (дата обращения: 23.01.2020).
4. Сайт НКО «Project for public spaces» [Электронный ресурс] Режим доступа: <https://www.pps.org/> (дата обращения: 23.06.2020).
5. РГИС [Электронный ресурс] Режим доступа: <https://rgis.spb.ru/mapui/> (дата обращения: 23.12.2020).



**Залуская Валерия Сергеевна**

Год рождения: 1998  
Университет ИТМО,  
факультет информационных технологий  
и программирования,  
студент группы №М41212,  
направление подготовки: 09.04.02 – Речевые  
информационные системы,  
e-mail: zvaleria.vz@yandex.ru



**Астапов Сергей Сергеевич**

Год рождения: 1988  
Университет ИТМО,  
факультет информационных технологий  
и программирования,  
ст.н.с., PhD,  
e-mail: sergeiastapov@gmail.com



**Кабаров Владимир Иосифович**

Год рождения: 1959  
Университет ИТМО,  
факультет информационных технологий  
и программирования,  
старший преподаватель,  
e-mail: vikabarov@itmo.ru



**Попов Дмитрий Владимирович**

Год рождения: 1979  
ООО «ЦРТ-Инновации»,  
в.н.с.,  
e-mail: ppdmitry@gmail.com

УДК 004.934

**ОСОБЕННОСТИ ЛЕКСИКИ В КОНТЕКСТЕ РЕШЕНИЯ ЗАДАЧИ  
ОПРЕДЕЛЕНИЯ ТОЧЕК СМЕНЫ ДИКТОРА**

**В.С. Залуская, С. Астапов, В.И. Кабаров  
Научный руководитель – в.н.с., Д.В. Попов**

Работа выполнена в рамках темы НИР №620172 «Определение структуры полилога с применением лексических и нелексических признаков речи нескольких дикторов» под руководством Астапова С.С.

### **Аннотация**

Рассмотрены основные особенности лексических данных телефонии в контексте решения задачи определения точки смены диктора. За основные факторы, влияющие на лексику, приняты формат диалога, расположение слов относительно границ реплик, тип перебивания, роль активного диктора в диалоге. Результаты экспериментов показывают, что в ролевой модели диалога существуют Ngram-ы, характерные для границ реплик дикторов.

### **Ключевые слова**

Речевая аналитика и понимание речи, Ngram-модель, точка смены диктора, разделение дикторов, диаризация, автоматическое распознавание речи.

### **Введение**

Процесс сегментирования аудио на промежутки, соответствующие речи одного диктора, и последующего указания идентификатора диктора называется диаризацией [1]. К одной из подзадач диаризации относится определение точек смены диктора, которые далее могут быть уточнены до четких границ реплик дикторов. Наибольшее распространение получили методы диаризации, основанные на анализе акустических характеристик речи и на пространственных признаках [2]. Однако данный подход игнорирует лексическую и семантическую составляющие реплик, что может быть покрыто NLU-методами (Natural Language Understanding, речевая аналитика и понимание речи) [3].

В текущей работе анализируется лексика в диалоге для дальнейшей оценки применимости NLU-методов в задаче определения точек смены диктора. Кроме того, применение NLU-методов необходимо в ситуациях, когда нет доступа к исходным аудио-файлам, а есть только текстовые транскрипции (например, в датасетах, основанных на субтитрах из фильмов [4]).

### **Описание данных**

В качестве данных для анализа используются двуканальные аудио телефонии из проприетарной базы данных ООО ЦРТ-Инновации, где в каждом из каналов располагается аудио с речью одного из дикторов. Кроме того, в данных изначально выделяется ролевая модель «Оператор-Клиент», где Оператор консультирует Клиента по возникшим у него проблемам и вопросам. Каждый из каналов обрабатывается моделью автоматического распознавания речи [5], в результате чего формируется корпус из текстовок, разделенных по сеансам обслуживания клиентов и включающих метки точек смены диктора. Всего в корпусе представлено 1196 сеансов обслуживания.

Первоначальная гипотеза заключается в том, что лексические данные содержат закономерности, обусловленные форматом диалога (телефонный, очный, переписка) и ролевой моделью диалога, которые позволяют достоверно оценивать вероятность присутствия точки смены диктора между подряд идущими словами. В частности, телефонный разговор имеет свою культуру общения, что выражается в паттернах обращения собеседников друг к другу. Также используемые данные содержат ярко выраженные роли дикторов, что дает основание полагать, что гипотеза может быть подтверждена.

В качестве метода разметки выбрано автоматическое распознавание речи. Во-первых, данный метод позволяет быстро получить транскрибированные и размеченные данные, по качеству сопоставимые с качеством модели распознавания речи. Во-вторых, перед началом исследования уже имелась проприетарная модель, разработанная ООО ЦРТ-Инновации. В-третьих, модель предоставляет информацию о границах слов на основе акустических признаков, что способствует удобному сведению двух каналов в единый текст с метками точек смены диктора. Однако модель автоматического

распознавания речи на выходе получает данные со специфическими ошибками, не свойственными ручной разметке, так как, например, выбор из слов, схожих по звучанию, осуществляется в зависимости от данных, на которых модель была обучена. Таким образом, полученные данные необходимо дополнительно обрабатывать для улучшения их качества.

При визуальном анализе телефонных диалогов была выделена одна из наиболее характерных особенностей данного формата речи – перебивания. Под перебиваниями понимаются временные фрагменты диалога, в которых слово одного диктора по своим границам наслаивается на слово другого диктора. В общем случае были выделены два типа перебиваний: пассивные и активные. Пассивные перебивания служат для поддержания диалога и выражения своей заинтересованности в реплике собеседника. При этом перебивание заключается в короткую реплику, после которой не происходит перехват инициативы. Активные перебивания, напротив, длительные и переходят в полноценную реплику с перехватом инициативы в разговоре. Таким образом, для обработки перебиваний использовались данные о временных границах слов, а также о длительностях реплик. Пассивные перебивания не несут информативного содержания, поэтому подлежат удалению, тогда как активные выделяются в отдельную реплику с выставлением метки точки смены диктора.

Таким образом, для анализа использовались данные телефонии в ролевой модели «Оператор-Клиент», распознанные моделью автоматического распознавания и очищенные от неинформативных перебиваний.

#### **Анализ языковых моделей на основе Ngram-моделей**

Для анализа лексики используется Ngram-ный подход: формируется языковая модель на основе Ngram – подпоследовательностей слов в речи. В качестве оценочного параметра языковой модели используется вероятность появления конкретной Ngram-ы в заданном контексте.

В работе используется несколько языковых моделей: модель, базирующаяся на вероятности того, что Ngram-а является окончанием реплики диктора, и модель, базирующаяся на вероятности того, что Ngram-а является началом реплики диктора. При этом также сохраняются метки принадлежности Ngram-ы конкретному диктору. Для анализа использовались Ngram-ы размерности от 3 до 5 включительно.

При анализе Ngram обеих языковых моделей без привязки к принадлежности к спикерам наиболее вероятные выражения на границах реплик соответствуют шаблонным фразам, свойственным официальному обращению, выражениям вежливости, репликам, направленным на запрос и предоставление информации.

В табл. 1 и табл. 2 указаны примеры начала и окончаний реплик соответственно.

Таблица 1

#### **Пять примеров начала реплик с указанием типа**

| Ngram                           | Вероятность | Тип                       |
|---------------------------------|-------------|---------------------------|
| Скажите как я могу              | 0.885       | Шаблонная фраза (скрипт)  |
| Спасибо оставайтесь пожалуйста  | 0.882       | Шаблонная фраза (скрипт)  |
| Скажите у вас остались          | 0.857       | Запрос информации         |
| Спасибо за звонок всего доброго | 0.78        | Вежливость                |
| Ну вот смотрите                 | 0.75        | Предоставление информации |

Таблица 2

**Пять примеров окончаний реплик с указанием типа**

| Ngram                            | Вероятность | Тип                      |
|----------------------------------|-------------|--------------------------|
| Всего вам доброго до свидания    | 1.0         | Вежливость               |
| У вас есть возможность подождать | 1.0         | Шаблонная фраза (скрипт) |
| Звонок всего доброго до свидания | 0.984       | Вежливость               |
| Я могу к вам обращаться          | 0.901       | Шаблонная фраза (скрипт) |
| Правильно я вас понимаю          | 0.791       | Запрос информации        |

Также на Ngram-ных моделях можно проследить влияние ролевой модели «Оператор-Клиент» на лексику. Так, Оператору свойственны шаблонные фразы, основанные на установленных компанией скриптах. При этом, так как каждый сеанс обслуживания носит консультационный характер, преимущественно используется вопросно-ответная форма диалога, в которой Оператору свойственны просьбы ожидания (во время решения запроса Клиента) и запрос уточнений, а Клиенту – запрос дополнительной информации и ответ на уточнения Оператора. Кроме того, Клиент выражается менее формально и шаблонно, что видно, например, по фразам благодарности. В табл. 3 и табл. 4 представлены примеры Ngram Оператора и Клиента соответственно с указанием типа.

Таким образом, можно сказать, что данные телефонии имеют свои лексические особенности, обусловленные форматом диалога и ролевой моделью, что позволяет выявить паттерны начала и окончания реплики.

Таблица 3

**Шесть примеров Ngram, соответствующих границам реплик Оператора**

| Ngram                         | Граница реплики | Вероятность | Тип                       |
|-------------------------------|-----------------|-------------|---------------------------|
| Всего вам доброго до свидания | Окончание       | 1.0         | Вежливость                |
| Так одну минуту               | Окончание       | 1.0         | Запрос ожидания           |
| Еще вопросы ко мне остались   | Окончание       | 1.0         | Запрос информации         |
| Так сейчас посмотрим          | Начало          | 1.0         | Предоставление информации |
| Скажите как я могу к          | Начало          | 0.878       | Шаблонная фраза (скрипт)  |
| Хорошо оставайтесь на линии   | Начало          | 0.875       | Запрос ожидания           |

Таблица 4

**Шесть примеров Ngram, соответствующих границам реплик Клиента**

| Ngram                 | Граница реплики | Вероятность | Тип                       |
|-----------------------|-----------------|-------------|---------------------------|
| Все спасибо большое   | Окончание       | 1.0         | Вежливость                |
| Ладно спасибо девушка | Окончание       | 1.0         | Вежливость                |
| Куда я могу позвонить | Окончание       | 1.0         | Запрос информации         |
| Так а скажите         | Начало          | 0.857       | Запрос информации         |
| Угу все спасибо       | Начало          | 0.833       | Вежливость                |
| Ну наверно            | Начало          | 0.75        | Предоставление информации |

### Результаты экспериментов

Помимо анализа языковых моделей реализованы Ngram-модели, определяющие на основе языковых моделей расположение точек смены диктора в текстовке речи. Реализованы модели на языковой модели в контексте начала реплики, на языковой модели в контексте окончания реплики и на использовании информации о начале и окончании реплики одновременно.

При этом у всех трех моделей метрика Recall имеет значение менее 0.5, что обусловлено большим количеством пропущенных точек смены диктора. Это объясняется тем, что более 50% точек окружены не типичными по лексике границами реплик, то есть относительно редко встречающимися в корпусе, что не позволяет модели выделить их из текста. Однако количество ложных срабатываний при этом мало, что отражается в метрике Precision, принимающей значения более 0.93 для всех трех моделей.

Использование гибридной модели, базирующейся на двух языковых моделях (для начала и окончаний реплик), не дало заметного прироста. Напротив, значение метрики Precision, использовавшейся как целевой, принимало свое максимальное значение при учете только окончаний реплик, тогда как добавление информации о начале реплик снижало ее значение.

### Заключение

В текущей работе проанализированы данные телефонии, выявлены особенности лексики для данного формата диалога и для ролевой модели «Оператор-Клиент», выделены зависимости лексики от начала и окончания реплики. Результаты показывают, что применение методов NLU для уточнения границ реплик дикторов применимо в описанном контексте.

### Литература

1. Anguera Miro, X. Speaker Diarization: A Review of Recent Research / Miro X.A., Bozonnet S., Evans N., Fredouille C., Friedland G., Vinyals O. // IEEE Transactions on Audio, Speech, and Language Processing. 2012. Vol. 20. No.2. P.356–370.
2. Астапов С.С., Лаврентьев А.В., Кабаров В.И. “Применение полигармонических моделей для определения пространственных признаков речи” // Труды всероссийской акустической конференции, Материалы III Всероссийской конференции. Изд. Федеральное государственное автономное образовательное учреждение высшего образования "Санкт-Петербургский политехнический университет Петра Великого". СПб. 2020. С. 350–355.
3. Anidjar O.H. A Thousand Words are Worth More Than One Recording: NLP Based Speaker Change Point Detection. / Or Haim Anidjar, et al. // arXiv preprint arXiv:2006.01206. 2020.
4. Tiedemann J. News from OPUS-A collection of multilingual parallel corpora with tools and interfaces / Tiedemann J. // In RANLP. 2009. P. 237–248.
5. Yu D. AUTOMATIC SPEECH RECOGNITION. / Yu D., Deng L. // Springer london limited. 2016.



**Землянова Екатерина Игоревна**

Год рождения: 1995  
Университет ИТМО,  
Институт дизайна и урбанистики,  
студент группы № С41801,  
направление подготовки: 27.04.07 – Научно-исследовательские  
технологии и экономика инноваций,  
e-mail: katya-zemlyanova95@yandex.ru



**Брагина Татьяна Владимировна**

Год рождения: 1997  
Университет ИТМО,  
Институт дизайна и урбанистики,  
студент группы № С41801,  
направление подготовки: 27.04.07 – Научно-исследовательские  
технологии и экономика инноваций,  
e-mail: bragina.arch@yandex.ru



**Лаушкина Анастасия Александровна**

Год рождения: 1996  
Университет ИТМО,  
Институт дизайна и урбанистики,  
студент группы № В42802,  
направление подготовки: 27.04.07 – Научно-исследовательские  
технологии и экономика инноваций,  
e-mail: nastasjalausckina@mail.ru



**Рослякова Светлана Витальевна**

Год рождения: 1988  
Университет ИТМО  
факультет инфокоммуникационных технологий,  
аспирант группы № 7772,  
направление подготовки: 09.06.01 – Информатика  
и вычислительная техника,  
e-mail: svetlana.roslyakova@itmo.ru



**Балаев Андрей Федорович**

Год рождения: 1981  
Университет ИТМО,  
Институт дизайна и урбанистики.  
к.т.н. доцент,  
e-mail: afbalaev@itmo.ru

УДК 004.89

## **ИССЛЕДОВАНИЕ ВЛИЯНИЯ РАЗЛИЧНЫХ СВЕТОВЫХ РЕЖИМОВ НА РАЗРЕШЕНИЕ КОНФЛИКТНОЙ СИТУАЦИИ В МНОГОПОЛЬЗОВАТЕЛЬСКОМ ПРОСТРАНСТВЕ**

**Е.И. Землянова, Т.В. Брагина, А.А. Лаушкина, С.В. Рослякова**  
**Научный руководитель – к.т.н., доцент А.Ф. Балаев**

Работа выполнена в рамках темы НИР №620185 «Разработка интеллектуальной многопользовательской адаптивной системы освещения для снижения конфликтного поведения и стрессовых ситуаций в образовательной среде».

### **Аннотация**

В статье рассмотрены тенденции развития интеллектуальных систем освещения, как они взаимодействуют с пользователем. Причины перехода на светодиодное освещение и их преимущества. Рассмотрено исследование IT-компании «КРОК» по применению адаптивной системы освещения у себя в офисе с целью повышения энергоэффективности и пользе такого освещения для человека.

### **Ключевые слова**

Адаптивное освещение, умные системы освещения, светодиоды, энергоэффективность, биодинамическое освещение.

Данная тема рассматривается на примере применения адаптивных систем освещения (АСО) в многопользовательском пространстве. В ходе анализа научных статей было выявлено, что большинство исследований в области систем управления освещением посвящено вопросам энергосбережения, повышения безопасности и комфорта.

Современные АСО непригодны для многопользовательского взаимодействия, поскольку выставленные ранее настройки не подходят другим пользователям одной системы по причине смены вида деятельности, времени суток, физического или эмоционального состояния человека, а предыдущие настройки освещения корректируются новым пользователем, делая их более недоступными.

Технологии АСО, которые позволяют пользователям управлять освещением с помощью своих гаджетов, таких как смартфоны, компьютеры или пульта дистанционного управления, становятся все более актуальнее в современном мире. Рынок интеллектуальных технологий быстро растет, и освещение является одним из основных перспектив на этом рынке.

Сегодня, в условиях экологического дисбаланса, все больше компаний начинают обращать внимание на экологическую составляющую своей деятельности для поддержки своего устойчивого развития. Переход на светодиодное освещение – это скорее необходимость, продиктованная требованиями времени, экономией и заботой об энергетических ресурсах.

Причиной растущей популярности является ряд преимуществ светодиодов перед остальными источниками света:

1. Долгий срок службы и низкое энергопотребление, как следствие экономическая выгода.
2. Высокая светоотдача. Практически всю получаемую энергию светодиод преобразует в свет.
3. Возможность выбора цветовой температуры:
  - от привычного теплого белого света лампы накаливания от 2700 до 3000К;

- до холодного белого или дневного света 6500К или 5000К.
- 4. RGB светодиоды - дают возможность выбрать подсветку любого из цветов спектра и менять ее с помощью специального контроллера по своему усмотрению.
- 5. Простота регулирования светового потока у светодиодных светильников
- 6. Компактность, гибкость и разнообразие модулей светодиодов.
- 7. Светодиоды устойчивы к механическим повреждениям [2].

Пользу от АСО на своем опыте проверила IT-компания «КРОК»: в 2013 году они внедрили в свой офис АСО с целью повышения энергоэффективности и уменьшения затрат на электроэнергию, а к 2017 с уверенностью говорили не только об энергоэффективности, но и о пользе адаптивного освещения, которое позволяет погружать человека в комфортную для него зону освещения, где автоматически регулируется температура света в зависимости от времени суток, подстраивается под циркадные ритмы человека, так как АСО благоприятно воздействует на биоритмы человека за счет изменения количества света и спектра светового воздействия в течение дня по аналогии с солнечным светом.

В результате проведенных исследований они получили экономию 45% на этаже площадью чуть более 1300 квадратных метров с учётом человеческого фактора (выборочного отключения автоматики для собственного комфорта). Данные регулярно сравниваются с показателями соседнего этажа, где такая система не установлена. Также данная система дает возможность забыть про выключатели и максимально использовать естественное освещение. Рабочие помещения оказываются равномерно освещены независимо от удаленности рабочего места от окна. В целом обеспечиваются комфортные условия работы, система работает незаметно и при этом экономит электроэнергию [4].

Также исследования показали, что люди, чьи офисы включали адаптивное освещение, на 15,2% лучше справляются со стрессом, чем те, что находятся в «широко освещенных» офисных помещениях [3].

Исследования Human Centric Lighting (HCL) показали, что динамические световые решения, имитирующие естественный ход дневного света, значительно улучшают настроение, концентрацию внимания, продуктивность, а также активность [1].

АСО создают комфортные условия для жизнедеятельности, сохранения здоровья и оптимальной работоспособности, а также способствуют повышению производительности труда и энергосбережению, что делает их актуальными для работодателей и общества в целом.

Где может эффективно работать АСО?

1. Учреждения здравоохранения: АСО подстраиваются под естественный ход дневного света и учитывает нужды пожилых пациентов – например, освещение для пациентов пожилого возраста должно быть ярче, чем для молодежи, или такое освещение может эффективно работать в реабилитационных центрах, где пациенты продолжительное время находятся в замкнутых пространствах и тем самым могут испытывать нехватку дневного света.

2. Школы: правильно выставленный режим АСО может помочь повысить сосредоточенность на выполнении задания. Например, при написании экзамена.

3. Офисы: применение АСО может помочь сократить затраты на электроэнергию так как система управляет освещением исходя из светового дня и необходимого уровня освещенности [3].

Таким образом, АСО благоприятно влияют на психоэмоциональное состояние человека, его работоспособность. Так как рабочий день можно разделить на несколько этапов работоспособности от погружения в рабочий процесс, устойчивой работоспособности до минимальной работоспособности в конце дня. Поэтому освещение должно создавать визуальный комфорт человеку и необходимый для работы

уровень освещенности. Например, освещение может помочь человеку вернуться из расслабленного послеобеденного состояния в сконцентрированное. Также снижается энергопотребление и затраты на электроэнергию.

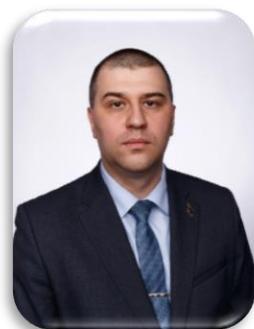
### Литература

1. Биологически и эмоционально эффективное освещение Human Centric Lighting [Электронный ресурс]. <https://www.ltcompany.com/ru/solutions/hcl/> (дата обращения: 22.01.2021).
2. Преимущества и недостатки светодиодных светильников, срок службы светодиодов [Электронный ресурс]. Режим доступа: <http://www.diy.ru/post/6240/> (дата обращения: 22.01.2021).
3. Умный свет в офисе экономит нам 40% затрат на электроэнергию [Электронный ресурс]. Режим доступа: <https://rb.ru/opinion/umnyj-svet-ekonomen/> (дата обращения: 22.01.2021).
4. Блог компании КРОК [Электронный ресурс]. Режим доступа: <https://habr.com/ru/company/croc/blog/178425/> (дата обращения: 22.01.2021).



**Исмагилова Аделина Фаритовна**

Университет ИТМО,  
факультет инфокоммуникационных технологий,  
студент группы №К42111,  
направление подготовки: 11.04.02 – Инфокоммуникационные  
технологии и системы связи,  
e-mail: is.adel.far@yandex.ru



**Басов Олег Олегович**

Университет ИТМО,  
факультет инфокоммуникационных технологий  
д.т.н., доцент, профессор,  
e-mail: oobasov@mail.ru

УДК 37.018.43

**АНАЛИЗ СУЩЕСТВУЮЩИХ  
СРЕДСТВ ОЦЕНИВАНИЯ УДОВЛЕТВОРЕННОСТИ  
СТУДЕНТОВ ЗАНЯТИЕМ В ДИСТАНЦИОННОМ ФОРМАТЕ  
ПОСРЕДСТВОМ ВИДЕОКОНФЕРЕНЦСВЯЗИ**

**А.Ф. Исмагилова**

**Научный руководитель – д.т.н., доцент, профессор О.О. Басов**

Работа выполнена в рамках темы НИР №620186 «Разработка средств оценивания уровня удовлетворенности студентов процессом дистанционного обучения посредством видеоконференцсвязи».

**Аннотация**

В данной работе представлена актуальность разработки программного средства и методы оценки уровня удовлетворенности студентов процессом дистанционного обучения. Определены понятия удовлетворенности студентов и качественного образования, на основе которых выполнен анализ существующих средств оценивания эмоционального состояния преподавателя и учащихся с целью повышения качества дистанционного обучения.

**Ключевые слова**

программное средство, видеоконференцсвязь, дистанционное обучение, удовлетворенность, психоэмоциональное состояние.

В настоящее время мир переживает период пандемии коронавирусной инфекции, в результате чего возникла необходимость перехода на дистанционное и смешанное обучение. В связи с этим возникла острая необходимость в создании новых методов оценки дистанционного образования [1, 2]. Подобные средства оценивания удовлетворенности студентов процессом дистанционного обучения создают

возможность повышения качества обучения путем создания оценочно-рекомендательных систем регулирования работы преподавателя.

Задача проекта заключается в создании программного обеспечения, которое позволит оценивать условия работы преподавателя, выявлять факторы влияния на эффективность управления учебным процессом и составить индивидуальные рекомендации преподавателю по повышению качества учебного процесса.

Прежде чем приступить к выполнению анализа средств оценивания, необходимо:

- определить компоненты, на основе которых будут определены понятия удовлетворенности студентов и качества образования;
- разработать формальный показатель удовлетворенности студентов процессом дистанционного обучения посредством видеоконференцсвязи;
- разработать алгоритм оценивания удовлетворенности студентов занятием;
- проанализировать средства оценивания удовлетворенности студентов дистанционным обучением.

Определение метрик для оценки степени удовлетворенности выполнялось методом опроса студентов, в результате которого было выделено, что больше всего на эмоциональное состояние влияют качество преподавания, время проведения занятия, комфортность местоположения во время занятия и факторы внешней среды. К подобным факторам можно отнести время года, суток, метеорологические изменения, ландшафт, звуковая зашумленность, освещенность и т.д. [3] Установив необходимые компоненты, было проведено анкетирование, в котором обучающимся следовало оценить уровень удовлетворенности учебным процессом по 5 балльной шкале (таблица).

Таблица

**Оценка степени удовлетворенности студентов**

| Метрика                   | % опрошенных, давших оценку |      |      |      |      |   | Степень удовлетворенности |
|---------------------------|-----------------------------|------|------|------|------|---|---------------------------|
|                           | 5                           | 4    | 3    | 2    | 1    | 0 |                           |
| Качество преподавания     | 0,5                         | 0,3  | 0,1  | 0,05 | 0,05 | 0 | 4,15                      |
| Время проведения занятия  | 0,3                         | 0,4  | 0,15 | 0,1  | 0,05 | 0 | 3,8                       |
| Комфортное местоположение | 0,3                         | 0,25 | 0,25 | 0,1  | 0,1  | 0 | 3,55                      |
| Погода                    | 0,25                        | 0,2  | 0,3  | 0,25 | 0    | 0 | 3,45                      |

Разработаем формулу удовлетворенности студентов и запишем ее в качестве коэффициента:

$$K_{\text{уд.ст.}} = \sum_{i=1}^n Y_i * K_i, \quad (1)$$

где  $K_{\text{уд.ст.}}$  – коэффициент «удовлетворенности студентов»;  $n$  – количество опрошенных студентов;  $i$  – номер опрошенного студента;  $Y_i$  – процент опрошенных студентов;  $K_i$  – Оценка по 5 б шкале.

Подставив результаты опроса в формулу 1, было выявлено, что коэффициент «удовлетворенности студентов» непостоянен и имеет прямую зависимость от качества образовательного процесса.

Под качественным образованием подразумевают организацию взаимодействия преподавателя и студентов, которая отражает степень соответствия учебного процесса

нормативным требованиям и удовлетворяет личностным и общественным запросам. Основные его понятия характеризуют достижение поставленных целей, имея определенный уровень квалификации: способность работать с информацией, степень овладения знаниями в предметной области, мотивация самообразования [4].

Разработаем формулу качественного образования:

$$K_{\text{кач.об.}} = \frac{\frac{\sum_{i=1}^k X_{iГ} * K_{iГ}}{K_{\Pi}}}{\frac{\sum_{i=1}^k Y_{iГ} * K_{iГ}}{K_{\Pi}}}, \quad (2)$$

где  $K_{\text{кач.об.}}$  – коэффициент «качественного образования»;  $X_{iГ}$  – среднее количество недовольных студентов в группе;  $K_{iГ}$  – количество студентов в группе;  $K_{\Pi}$  – количество студентов в потоке;  $Y_{iГ}$  – максимально возможное количество недовольных студентов в группе (определяется эмпирическим путем).

Определив формальные показатели удовлетворенности студентов процессом дистанционного обучения посредством видеоконференцсвязи, возникает понимание связи учебного процесса с эмоциональным состоянием обучающегося и преподавателя. На основе полученных данных разработаем алгоритм работы системы определения уровня удовлетворенности студентов.

На рис. 1 показано, что дистанционное образование вкладывает свои условия определения удовлетворенности, поскольку занятия переходят в видеоконференцсвязь. Таким образом, необходимо разделять аудио-, видео- и текстовые потоки занятия и на их основе определять не только эмоциональное состояние преподавателя и студентов, но и качество учебного материала (оформление презентаций согласно требованиям), временные параметры и вербальные характеристики.

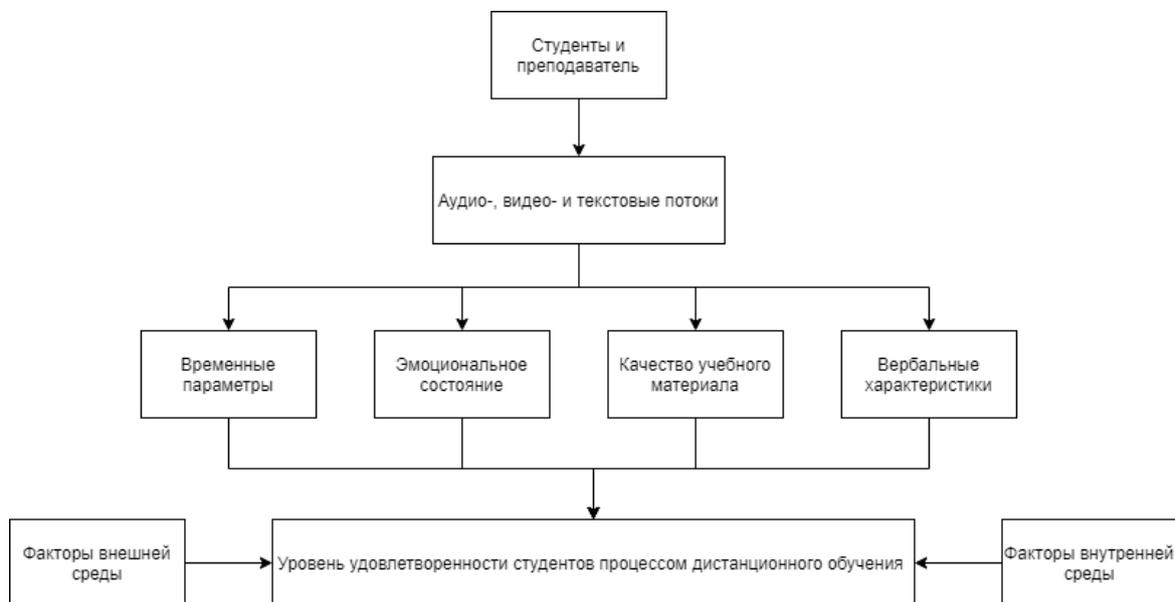


Рис. 1. Алгоритм работы системы по определению удовлетворенности студентов процессом дистанционного обучения

Отталкиваясь от алгоритма, возникает необходимость анализа существующих аналогов по определению психоэмоционального состояния. В настоящее время существуют программно-аппаратные средства для распознавания эмоций, такие как FaceReader [5], eMotion Software [6], Affectiva Affdex [7], NuraLogix [8] и т.д.

FaceReader[5] позволяет определить семь базовых эмоций, пол, возраст и создавать модель лица по контрольным точкам. Данный сервис имеет точность распознавания до 80% (рис. 2).



Рис. 2. Интерфейс сервиса FaceReader

eMotion Software [6] позволяет определить шесть базовых эмоций (гнев, печаль, страх, удивление, отвращение и счастье) и создавать 3D модели лица с помощью выявления 12 ключевых областей (рис. 3).

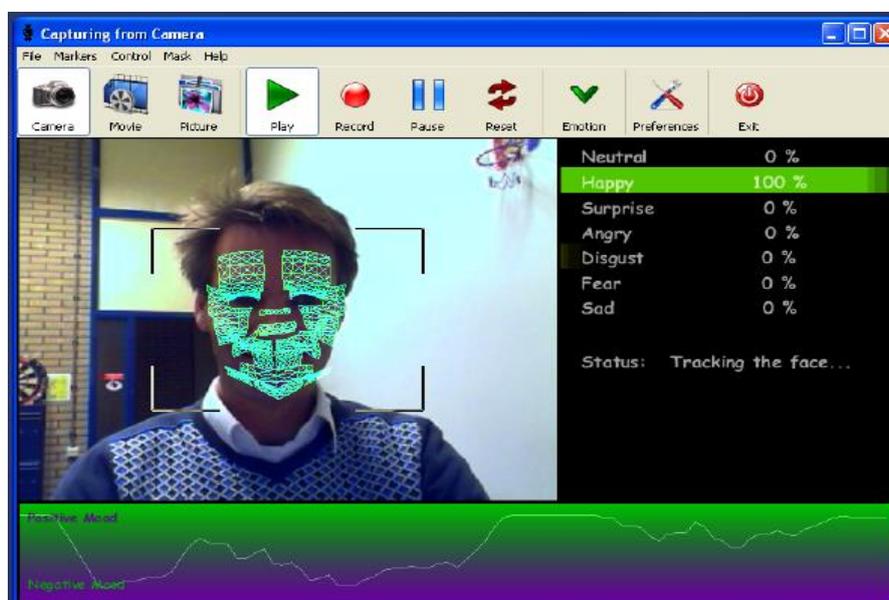


Рис. 3. Интерфейс системы eMotion Software

Affectiva Affdex [7] позволяет распознавать 7 базовых эмоций с использованием камеры смартфона: выполняется запись изображения лица, вейвлет-преобразование и сравнение с образцами на основе SIFT-дескрипторов (рис. 4).



Рис. 4. Результат работы комплекса Affectiva Affdex

NuraLogix [8] позволяет распознавать эмоции на основе цвета подкожного узора кровеносных сосудов лица. Для определения цвета пикселя лица (изменение кровотока) используются трансдермальные оптические изображения (рис. 5).

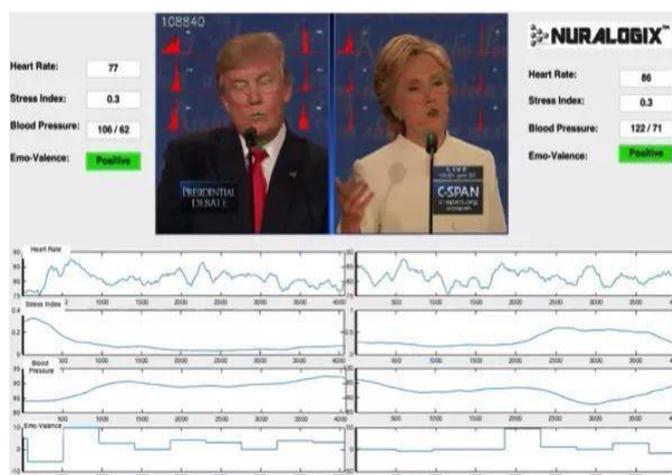


Рис. 5. Результат работы NuraLogix

В ходе исследования было выявлено, что существующие средства распознавания эмоций не позволяют анализировать психоэмоциональное состояние большого количества людей одновременно и разрабатывать рекомендации по полученным результатам.

В дальнейшей работе планируется разработка программного средства, которое позволит определять удовлетворенность студентов и создавать оценочно-рекомендательную системы регулирования работы преподавателя на основе многомодальных данных занятия, проводимого в режиме видеоконференцсвязи.

## Литература

1. Маслова М.А., Лагуткина Т.В. Анализ и выявление положительных и отрицательных сторон внедрения дистанционного обучения // Научный результат. Информационные технологии. Т.5, №2, 2020. [http://rrinformation.ru/media/information/2020/2/%D0%98%D0%A2\\_8.pdf](http://rrinformation.ru/media/information/2020/2/%D0%98%D0%A2_8.pdf). (дата обр. 01.01.2020).

2. Chris Panagiotakopoulos, Antonis Lionarakis, Michalis Xeno: Open and Distance Learning: Tools of Information and Communication Technologies for Effective Learning (2003).
3. Станкевич Е.Ю. К вопросу оценки качества образования // Гуманитарные научные исследования. 2013. №1. [Электрон. ресурс] URL: <http://human.snauka.ru/2013/01/2215> (дата обращения: 16.12.2020).
4. Исмагилова А.Ф., Дудина Д.С., Алейников С.А. Разработка средств оценки уровня удовлетворенности студентов процессом дистанционного обучения посредством видеоконференцсвязи // Научный результат. Информационные технологии – 2020. Т. 5. № 4. С. 44-52.
5. M.J. den Uyl, H. van Kuilenburg. The FaceReader: Online Facial Expression Recognition (2005).
6. Shunji Mitsuyoshi. Emotion recognizing method, sensibility creating method, device, and software (2008).
7. McDuff D., Mahmoud A., Mavadati M., Amr M., Turcot J., Kaliouby R.E. AFFDEX SDK: A Cross-Platform Real-Time Multi-Face Expression Recognition Toolkit. CHI EA '16: Proceedings of the 2016 CHI Conference Extended Abstracts on Human Factors in Computing Systems. pp. 3723–3726. 2016.
8. Lee K., Kabakov E., Levy P. System and method for camera-based heart rate tracking (2018).



**Капранова Влада Олеговна**

Год рождения: 1998  
Университет ИТМО,  
факультет информационных технологий  
и программирования,  
студент группы №М41211,  
направление подготовки: 09.04.02 – Информационные  
системы и технологии,  
e-mail: vlada-kapranova@mail.ru



**Астапов Сергей Сергеевич**

Год рождения: 1988  
Университет ИТМО,  
факультет информационных технологий  
и программирования,  
ст.н.с., PhD,  
e-mail: sergeiastapov@gmail.com



**Кабаров Владимир Иосифович**

Год рождения: 1959  
Университет ИТМО,  
факультет информационных технологий  
и программирования,  
старший преподаватель,  
e-mail: vikabarov@itmo.ru



**Попов Дмитрий Владимирович**

Год рождения: 1979  
ООО «ЦРТ-Инновации»,  
в.н.с.,  
e-mail: ppdmitry@gmail.com

УДК 004.934

**СРАВНЕНИЕ МЕТОДОВ ВЫДЕЛЕНИЯ РЕПЛИК ЦЕЛЕВОГО  
ДИКТОРА В ЗАПИСИ РЕЧИ НЕСКОЛЬКИХ ДИКТОРОВ**

**В.О. Капранова, С.С. Астапов, В.И. Кабаров**

**Научный руководитель – в.н.с., ООО ЦРТ-Инновации Д.В. Попов**

Работа выполнена в рамках темы НИР №620172 «Определение структуры полилога с применением лексических и нелексических признаков речи нескольких дикторов».

**Аннотация**

В работе рассмотрены методы VoiceFilter и GSS с целью определения преимуществ и недостатков методов для решения задачи выделения реплик целевого диктора из

аудиозаписи, а также проведено их сравнение. По результатам сравнения определены преимущества методов VoiceFilter и GSS и на основе анализа был выбран метод для работы с базами данных, аналогичных базе CHiME-5.

#### **Ключевые слова**

Диаризация, целевой диктор, VoiceFilter, Guided Source Separation, GSS, The CHiME-5 Challenge.

Внимание человека давно привлекают методы, которые могли бы автоматизировать разного рода рутинные задачи. Особый интерес представляет распознавание речи на аудиозаписях и их разметка как ресурсоёмкая и монотонная для человека работа. Автоматические методы распознавания речи существенно ускорят и упростят решение такой задачи. Поэтому для анализа речи в рамках диалога требуются автоматические средства разметки реплик для каждого диктора на аудиозаписи. Проблема решается с помощью диаризации, которая является процессом определения временных границ высказывания для целевого диктора в разговоре. Надо отметить, что эта задача не решена, так как всё усложняется тем, что в речи могут присутствовать перебивания (наложение речи нескольких дикторов на одном временном отрезке). Для решения разделения речи на реплики каждого из участников существует два класса методов: на основе сквозной нейронной диаризации (EEND) и с использованием предварительно вычисленных характеристик интересующего говорящего, чтобы привлечь внимание системы только к его речи (Target-Speaker ASR, Speaker Beam, Voice Filter [1], GSS [2]). В данной работе рассмотрены два из них – VoiceFilter и GSS как наиболее применимые и достаточно новые методы. Алгоритм Speaker Beam архитектурно похож на VoiceFilter, поэтому в работе он не рассматривался.

Сравнение методов производится с целью выявить метод, наиболее подходящий для работы с базой данных The CHiME-5 Challenge [2]. Данные имеют следующую специфику. Количество участников всегда фиксировано. Запись ведется при помощи шести 4-микрофонных массивов Microsoft Kinect в трёх зонах квартиры. Наличие разных зон записи на практике приводит к появлению шумов специфического характера на фоне. Общее количество записей равняется двадцати, из них шестнадцать записей применялись для обучения методов, две для дообучения и две для оценки качества. Всего доступно около 40 часов аудиозаписей. Такой ограниченный объём данных может существенно ухудшить результаты работы методов, для обучения которых требуются большое количество данных (порядка 1000 часов). В табл. 1 представлено процентное соотношение количества одновременно говорящих людей на аудиозаписях. Значение «0» соответствует ситуации, когда все участники разговора молчат.

Таблица 1

**Количество одновременно говорящих людей на аудиозаписи**

| Количество говорящих  | 0     | 1     | 2     | 3    | 4    |
|-----------------------|-------|-------|-------|------|------|
| Записи для дообучения | 24%   | 54.3% | 17.7% | 3.5% | 0.5% |
| Записи для оценки     | 33.5% | 51.5% | 12%   | 2.5% | 0.5% |

Для начала обратимся к системе VoiceFilter [1]. Его схема представлена на рис. 1.

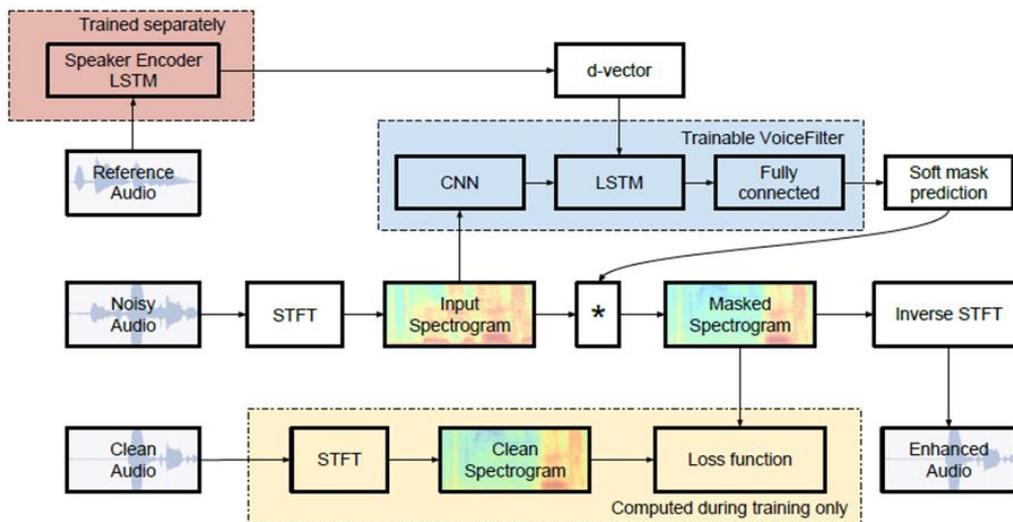


Рис. 1. Архитектура VoiceFilter [1]

Система состоит из трёх блоков: блок вычисления голосовой модели целевого диктора, состоящий из нейронной сети с долгой краткосрочной памятью, блок VoiceFilter и блок обратной связи. На выходе первой нейронной сети получается d-вектор (вычисленная голосовая модель), который подаётся на вход VoiceFilter. Внутри VoiceFilter состоит из 8 слоёв свёрточной нейронной сети, одного слоя Long short-term memory (LSTM) и двух полносвязных слоёв. На вход принимается вектор  $d$  и спектрограмма входного сигнала, вычисленная из входной зашумлённой записи. Выходом блока является мягкая маска со значениями от 0 до 1 и имеющая размер, равный размеру вычисляемой спектрограммы. Эта маска применяется к спектрограмме, вычисленной из зашумлённой аудиозаписи, и после применяется обратное преобразование Фурье для получения конечного результата.

Теперь перейдём к методу Guided Source Separation (GSS) [2]. Этот метод был написан в рамках соревнования CHiME5 [2]. Его схема приведена на рис. 2.

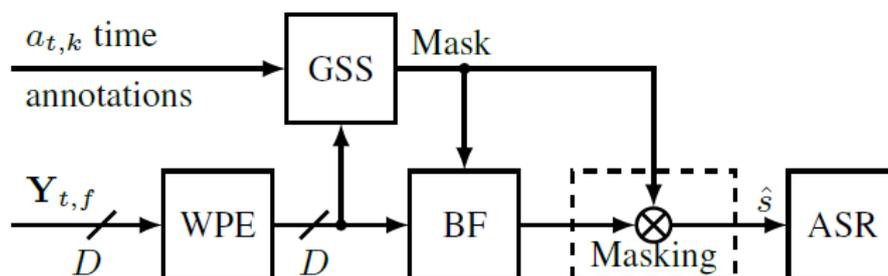


Рис. 2. Архитектура системы GSS [2]

Метод состоит из 5 блоков:

1. Многоканальная дераверберация на основе принципа взвешенной ошибки прогнозирования Weighted Prediction Error (WPE) – предварительный этап обработки сигнала.

2. Блок GSS выводит частотно-временные маски, из которых оцениваются пространственные ковариационные матрицы, и из этих матриц вычисляются коэффициенты для бимформинга (статистически оптимального формирователя диаграммы направленности с минимальной дисперсией без искажений Minimum Variance Distortionless Response (MVDR)).

3. Блок формирования направленного сигнала (Beamforming – BF) [3].

4. Спектральное маскирование (Masking) применяется для улучшения характеристик разделения при низком пространственном разнесении динамиков [4].

5. Блок автоматического распознавания речи (Automatic Speech Recognition – ASR).

На вход системы подаются многоканальный сигнал и временные аннотации. Данные временные аннотации указывают, когда конкретный говорящий активен (0 – неактивен, 1 – активен). Они представлены на рис. 3. Дополнительно вводится класс шума, который активен всегда. Необходимо учитывать контекст вокруг фразы, чтобы убедиться, что это целевой говорящий, а не класс шума. Если вокруг фразы достаточно нулей, то это целевой говорящий.

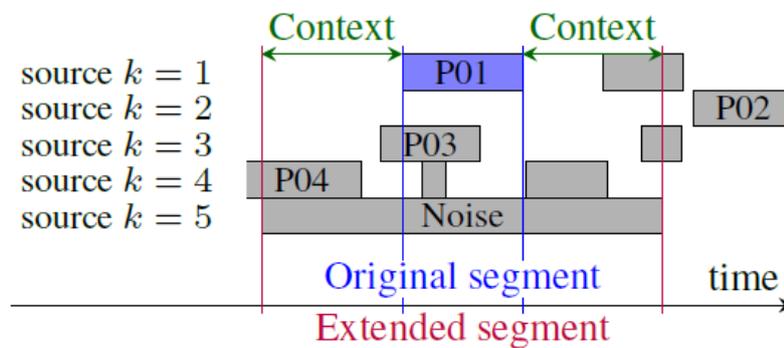


Рис. 3. Графическое представление временных аннотаций [2]

Сравнение двух рассмотренных выше методов приводится в табл. 2.

Таблица 2

Сравнение методов VoiceFilter и GSS

|              | Многоканальность | Работа на сильно зашумлённых данных | Улучшение аудиозаписей | Использование априорной информации | Возможность обобщения | Влияние синхронизации каналов | Время обучения метода |
|--------------|------------------|-------------------------------------|------------------------|------------------------------------|-----------------------|-------------------------------|-----------------------|
| Voice Filter | +                | -                                   | +                      | +                                  | +                     | -                             | 20 часов              |
| GSS          | +                | +                                   | -                      | +                                  | -                     | +                             | ~ 40 часов            |

Важным свойством является то, что оба метода работают с многоканальными записями. Также оба метода используют дополнительную априорную информацию (временные аннотации для GSS и опорные аудиозаписи для VoiceFilter).

Из преимуществ у VoiceFilter есть возможность дополнительного улучшения качества обрабатываемых аудиозаписей и возможность обобщения от одного набора данных к другому. Однако этому методу требуются большие объёмы данных для обучения.

Благодаря дополнительному спектральному маскированию GSS может работать на сильно зашумлённых данных, где не справляются другие методы, но на его работу сильно влияет синхронизация каналов, что надо отдельно контролировать.

VoiceFilter – более быстрый метод (для оценки времени взят примерно одинаковый объём тренировочных данных).

Из результатов сравнения видно, что VoiceFilter больше подходит для ситуаций, когда есть большие объёмы похожих данных (например, внутри одной сферы), а также со средней и слабой шумовой обстановкой: на судебных заседаниях, в относительно тихих условиях медицинских кабинетов, так как преимущество VoiceFilter перед GSS – большая скорость обучения и обобщение наборов данных. GSS же лучше справится с шумной обстановкой во время совещания и для контроля работы кассиров в магазинах, а также других людей, работающих в сфере обслуживания, что обусловлено преимуществом работы метода для сильно зашумлённых данных.

Суммируя вышесказанное, можно сделать вывод, что для работы с базой данных CHiME-5 лучше использовать метод GSS, так как он не требует больших объёмов данных, как в случае с VoiceFilter, а также нечувствителен к шумовой обстановке в аудиозаписи.

В ходе работы были сделаны следующие выводы. VoiceFilter быстрее, а GSS хорошо работает в очень шумных условиях. Общим недостатком таких методов является использование заранее известной информации, которая в реальных ситуациях не всегда доступна. При этом для работы с базой данных CHiME-5, а также с другими аналогичными по шумовой обстановке и объёму данных базами более применим GSS.

## Литература

1. Wang Q., Muckenhirn H., Wilson K., Sridhar P., Wu Z., Hershey J., Saurous R. A., Weiss R. J., Jia Y., Moreno I. L. “Voice-Filter: Targeted voice separation by speaker-conditioned spectrogram masking” (eds) Proc. Interspeech. 2019. P. 2728–2732.
2. Boeddecker C., Heitkaemper J., Schmalenstroer J., Drude L., Heymann J., Haeb-Umbach R. “Front-End Processing for the CHiME-5 Dinner Party Scenario” // CHiME Workshop. 2018. P. 35–40.
3. Astapov S., Lavrentyev A., Shuranov E. “Far Field Speech Enhancement at Low SNR in Presence of Nonstationary Noise Based on Spectral Masking and MVDR Beamforming”. In: Karpov A., Jokisch O., Potapova R. // Speech and Computer. SPECOM 2018. Lecture Notes in Computer Science. Vol 11096. 2018. P. 21–31.
4. Астапов С.С., Лаврентьев А.В., Кабаров В.И. “Применение полигармонических моделей для определения пространственных признаков речи” // Труды всероссийской акустической конференции, Материалы III Всероссийской конференции. Изд. Федеральное государственное автономное образовательное учреждение высшего образования “Санкт-Петербургский политехнический университет Петра Великого”, СПб, 2020. С. 350–355.



**Кирилова Светлана Анатольевна**  
Университет ИТМО,  
факультет инфокоммуникационных технологий,  
студент группы №К41103с,  
направление подготовки: 11.04.02 –Инфокоммуникационные  
технологии и системы связи,  
e-mail: kirilova.svetl.an@yandex.ru



**Осипов Никита Алексеевич**  
Университет ИТМО,  
факультет инфокоммуникационных технологий,  
к.т.н., доцент,  
e-mail: nikita@ifmo.spb.ru

УДК 004.04

## **ПРИМЕНЕНИЕ НЕЙРОННЫХ СЕТЕЙ В ПОСТРОЕНИИ ИНДИВИДУАЛЬНОЙ ТРАЕКТОРИИ ОБУЧЕНИЯ**

**С.А. Кирилова**

**Научный руководитель – к.т.н., доцент Н.А. Осипов**

### **Аннотация**

Существует множество исследований, охватывающих сферу построения индивидуальных траекторий обучения с использованием нейронных сетей. В работе рассматриваются сферы применения нейронных сетей, а также проблемы и перспективы применения нейронных сетей для построения индивидуальной траектории обучения.

### **Ключевые слова**

Нейронные сети, индивидуальная траектория обучение, рекуррентные нейронные сети.

На сегодняшний день в мире основным является образовательный процесс фронтальной формы, где занятия проходят в классах с определенным количеством людей, а подаваемая информация одинакова для каждого. Однако очевидным является факт того, что при индивидуальном обучении достигается наибольшая эффективность и усвояемость материала, поскольку учитываются способности каждого ученика, которые среди класса или группы могут варьироваться достаточно широко в связи с генетическими особенностями или пережитым опытом. Помимо того, что существует формат обучения, в котором внимание уделяется лишь одному человеку, преимущество такого подхода заметнее всего не при общем плане обучения, а скорее при индивидуально построенном пути, состоящем из тем и разделов, которые больше всего нуждаются в укреплении теоретически и практически. Подобный путь называется «Индивидуальная образовательная траектория».

Актуальность темы обосновывается постоянным ростом требований к качеству обучения в современных условиях в целом, и в частности, в условиях дистанционного взаимодействия между людьми. Очевидно, что наличие индивидуальной траектории обучения значительно увеличит уровень образования населения в целом и выведет на

новую ступень развитие сервисов самообразования, но выделение частного репетитора для каждого человека невозможно в связи с трудоемкостью процесса определения уровня способностей и знаний каждого. Одним из актуальных инструментов для построения точных траекторий обучения является машинное обучение и нейронные сети [1]. Целью данной работы является обзор и анализ научных исследований и разработок за последние пять лет, направленных на продвижение использования нейронных сетей для получения наиболее подходящих и удовлетворяющих пользователя индивидуальных траекторий обучения.

Так как построение индивидуальных траекторий является многоэтапным процессом, в результате анализа было выявлены те сферы, которые могут быть автоматизированы за счет применения нейронных сетей. В данный момент ведутся исследования в следующих областях:

- разработка множества интеллектуальных экспертных систем для оценки трудностей и классификации тематики задач. Для классификации задач применяются алгоритмы обработки естественного языка, такие как скрытое распределение Дирихле и неотрицательная матричная факторизация;

- нейронные сети используются в качестве инструмента моделирования приложений для прогнозирования успеваемости студентов в зависимости от различных факторов;

- разработка инструмента на основе искусственного интеллекта, который учитывает ИНС (искусственная нейронная сеть) как один из методов для минимизации дезориентации учебного поведения и перегрузки когнитивных проблем среди студентов;

- использование глубокой нейронной сети для оценки приобретения знаний и навыков учащимися. Модуль предназначен для формирования индивидуальной траектории обучения по средствам занятий, проходящих в электронном виде;

- в попытке компенсировать недостаток репетиторов, выдвигается идея использования машинного обучения для ассистирования учеников с целью рекомендации набора задач и материалов с учетом сильных сторон отдельных учащихся в каждой категории и их предпочтения. Данная система рекомендации (RC – recommendation system) траектории обучения основывается на диаграммах способностей учащихся, построенных при помощи рекуррентной нейронной сети.

Все перечисленные компоненты используются для построения траектории обучения на разных этапах при использовании нейронных сетей.

Однако, несмотря на множество исследований, некоторые проблемы непосредственно остаются остановочным фактором для широкого использования нейронных сетей при построении индивидуальных траекторий обучения.

Касаясь трудностей, с которыми сталкиваются исследователи в области применения нейронных сетей при разработке индивидуальных траекторий обучения, в первую очередь следует упомянуть недостаток данных для обучения любых видов ИНС. Требуется разработка методологий и техник по увеличению данных или же увеличения точности работы нейронных сетей при малом количестве исходной информации. Также существуют исследования, предлагающие решение данной проблемы за счет применения нейронных сетей. Например, в работе [2] предлагается новый метод увеличения данных – асинхронное расширение данных посредством предварительной категоризации (Asynchronous Data Augmentation through Pre-Categorization, ADAPC), который решает эту проблему.

Следующей проблемой является ограниченное использование ИНС в области образовательных исследований вследствие сложности моделирования из-за недостатка исходных данных [3]. Данный факт может быть связано с возрастающей сложностью моделируемой сети, сложностью моделируемой системы ИНС для предоставления

подходящего объяснения (природа черного ящика), склонностью к чрезмерной подгонке и временем, необходимым для обучения нейронной сети.

Также важно отметить, что рекомендательные системы, используемые для ассистирования учеников, сталкиваются с двумя основными проблемами: «Cold Start» и «Sparsity» [4] (переводя дословно – «Холодный старт» и «Разреженность» данных. «Разреженность» данных означает сложность поиска достаточно надежных похожих пользователей, поскольку в целом активные пользователи оценивают лишь небольшую часть элементов; «Холодный старт» означает сложность выработки точных рекомендаций для новых пользователей, которые оценили лишь небольшое количество элементов). Ряд исследователей предлагают решение этой проблемы с использованием глубокого обучения.

Приоритетом современного образования является адаптация методов и темпов передачи знаний и навыков в зависимости от индивидуальных предрасположенностей каждого отдельного ученика. С целью разделения процесса обучения на этапы, оценивания уровня владения тем или иным навыком и предположения о наиболее подходящих заданиях в данный момент при построении индивидуальных траекторий обучения широко используется ИНС. Нейронные сети применяются для классификации задач, оценки знаний, а также в рекомендательных системах.

Для реализации указанных функций применяются рекуррентные ИНС метода обратного распространения ошибки, а также глубинные нейронные сети, в частности сети долгой краткосрочной памяти. В целом ИНС может выполнять нейронную подгонку и прогнозирование, а также способна теоретически классифицировать любые данные с произвольной точностью. Однако существует лишь ограниченное количество исследований и инструментов для прогнозирования успеваемости учащихся, особенно в контексте интеллектуального образования.

Основная трудность использования нейронных сетей – недостаток данных для обучения любых видов ИНС. Требуется разработка методологий и техник по увеличению данных или же увеличению точности работы нейронных сетей при малом количестве исходной информации. Проблема, с которыми сталкивается среда электронного обучения в системах рекомендаций на основе контента (CB) и коллективного фильтра (CF), можно назвать «холодный старт» и «разреженность» данных, которые объединены недостатком данных об активности пользователей.

Будущие исследования должны ставить задачей разрешение указанных проблем путем применения методов глубинно обучения нейронных сетей, а также рассматривать построения моделей с увеличенным числом параметров, учитываемых ИНС, с целью достижения максимальной точности предсказываемых результатов.

## Литература

1. Filippov D.M. et al. Modeling the construction of an individual learning path //Journal of Physics: Conference Series. IOP Publishing. 2020. Т. 1679. №. 2. С. 022094.
2. Cader A. The Potential for the Use of Deep Neural Networks in e-Learning Student Evaluation with New Data Augmentation Method //International Conference on Artificial Intelligence in Education. Springer. Cham. 2020. С. 37-42.
3. Klačnja-Milićević A. et al. E-learning systems: Intelligent techniques for personalization. Springer, 2016. Т. 112.
4. Srivastav G., Kant S. Review on e-Learning Environment Development and context aware recommendation systems using Deep Learning //2019 3rd International Conference on Recent Developments in Control, Automation & Power Engineering (RDCAPE). IEEE, 2019. С. 615-621.



**Кшенин Александр Дмитриевич**

Год рождения: 1997

Университет ИТМО,

факультет цифровых трансформаций,

студент группы № С42131,

направление подготовки: 09.04.02 – Информационные

системы и технологии,

e-mail: Ksh.Al.Dm@gmail.com



**Ковальчук Сергей Валерьевич**

Год рождения: 1984

Университет ИТМО,

факультет цифровых трансформаций,

к.т.н., доцент,

e-mail: kovalchuk@itmo.ru

УДК 004.942

## **ИДЕНТИФИКАЦИЯ И ИССЛЕДОВАНИЕ ИМИТАЦИОННЫХ МОДЕЛЕЙ БИЗНЕС-ПРОЦЕССОВ В КРУПНОЙ КОМПАНИИ ПО ДАННЫМ КОРПОРАТИВНЫХ ИНФОРМАЦИОННЫХ СИСТЕМ**

**А.Д. Кшенин**

**Научный руководитель – к.т.н., доцент С.В. Ковальчук**

### **Аннотация**

В работе производится разработка программы для мультиагентного имитационного моделирования бизнес-процессов, внедряющего различные связи и взаимодействия сотрудников в исследуемые модели бизнес-процессов. Компания представляется в виде мультиплекс-сети её сотрудников, что позволяет использовать информацию из неё для моделирования взаимодействия сотрудников при выполнении бизнес-процессов, а также исследовать сопутствующую динамику слоёв сети.

### **Ключевые слова**

Моделирование бизнес-процессов, имитационное моделирование, мультиагентное моделирование, мультиплекс сети, корпоративные информационные системы.

Моделирование бизнес-процессов (БП) ставит задачу построения исполняемых цифровых представлений протекающих в компании процессов, доступных для дальнейшего анализа и оптимизации. Данное направление является широко востребованным в настоящее время, существует множество решений, продуктов и реализаций, позволяющих получить структуру БП компании в виде, в котором они действительно исполняются, провести моделирование БП на длительный срок, с различными параметрами и нагрузкой на систему. Однако важную роль при выполнении БП играют сотрудники компании, их связи и взаимодействия, что зачастую упускается в существующих моделях, прежде всего ориентированных на структурные связи между БП. При том, что конкретные типы взаимодействий, связей

между сотрудниками могут различаться в зависимости от конкретной задачи, в общем виде они представимы динамической многослойной мультиплекс-сетью (МС). Внедрение социальных взаимодействий в моделирование различных элементов функционирования компании, различных её БП является новой актуальной задачей [1].

В данной работе производится разработка программы, реализующей мультиагентное имитационное моделирование протекающих в компании БП, использующее информацию из МС сотрудников. Сотрудники представляют собой действующих агентов моделирования, взаимодействующих, принимающих решения и напрямую влияющих на исполнение БП [2]. Они совершают действия, основываясь на своих характеристиках, связях в МС и реализуемых БП, происходит интеграция моделирования БП и динамики МС. Обратная задача, решаемая подобным подходом, заключается в том, что связи между БП в цепочках трансформируются при взаимодействии сотрудников в эволюционирующие связи между ними, что позволяет исследовать динамику сети, изменение различных, в том числе мультиплекс, характеристик вершин [3]. Цепочки БП (ЦБП) состоят из последовательно исполняемых связанных БП и определяются посредством обработки связей между БП в информационной системе (ИС) компании.

Предлагаемая в работе программа реализована на Python 3 (версии 3.7), в основе работы лежит фреймворк процессо-ориентированной дискретно-событийной системы моделирования SimPy (версии 4.0.1), а также использует библиотеки SALib (версии 1.3.11), ru3plex (версии 0.75), NetworkX (версии 2.5) и igraph (версии 0.8.3).

Структура программы включает:

- основу имитационного моделирования, включающую создание и запуск среды (environment) фреймворка SimPy;
- объекты данных – Python классы БП, ЦБП, реализаций БП и ЦБП;
- агентов моделирования – классы менеджера и сотрудников, содержащие метод запуска run – процесс (process), определяемый в фреймворке SimPy;
- МС компании, связывающую сотрудников, которая представлена в виде графа NetworkX;
- визуализатор динамики МС, использующий методы из библиотеки ru3plex;
- модуль отслеживания, агрегации и обработки статистики моделирования;
- генератор синтетических данных: БП, ЦБП, МС компании;
- интерфейс импорта данных, трансформации в формат для моделирования;
- модуль анализа чувствительности, использующий методы из библиотеки SALib;
- набор интерактивных блокнотов (.ipynb) для запуска экспериментов с моделью в Jupyter Notebook.

На вход программе подаётся информация о связях между БП в цепочках, связях между сотрудниками в МС, связях сотрудник–БП, т.е. компетенциях – информации об исполняемых сотрудником БП, дополненных временем или параметрами времени исполнения. После инициализации модели производится введение пользовательских ограничений, например, на исследуемый отдел компании, генерация параметров распределений для компетенций сотрудников (при отсутствии в данных), построение ЦБП из связей, с учётом введённых ограничений.

В модели заданы пуассоновские точечные процессы генерации реализаций ЦБП с экспоненциальным временем между появлением. Менеджер – агент моделирования, представляющий корпоративную ИС, распределяет новые БП между исполнителями, обладающими соответствующими компетенциями. Стратегия распределения является пользовательской функцией и может быть, например, выдачей реализации БП случайному сотруднику или сотруднику с наименьшей нагрузкой. При моделировании

работы над реализацией БП сотрудники тратят случайное время, генерируемое из параметров, содержащихся в компетенциях, переводят реализацию БП между состояниями Открытый, Активный, Завершённый, Закрытый. Выполнение реализации БП соответствует переводу её из состояния Активный в Завершённый, состояние Закрытый, при наличии нескольких сотрудников, требует выполнения работы от каждого из них. Т.к. БП в цепочке выполняются последовательно, то перевод из Открытого в Активный производится при завершении всех работ над предыдущим БП в цепочке.

В текущей реализации МС рассматриваются слои: территориальной структуры отделов компании (пространственные связи, Geography), её иерархической организации (руководители–подчинённые, Hierarchy), общих компетенций сотрудников (сходство по мере Жаккара множеств БП, выполняемых сотрудниками, Competencies), их рабочих контактов (фактов взаимодействия при выполнении БП, Contacts), дополненные слоем выполнения БП (связи, возникающие при выполнении БП и передаче цепочки, Processes). На рис. 1 представлена визуализация МС синтетической (сгенерированной) компании. Визуализация соответствует отдельному моменту времени моделирования выполнения БП. Связи в слоях компетенций и контактов исчезают со временем.

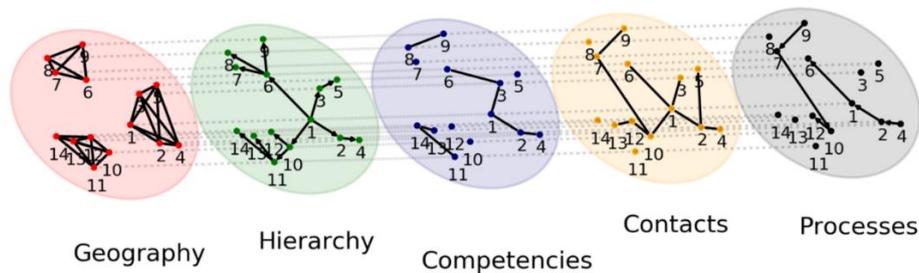


Рис. 1. Мультиплекс сеть сгенерированной компании

Для мониторинга эффективности реализации БП в программе производится логирование перевода БП между состояниями, времён и характеристик исполнения, а также вычисление отклика моделирования, включающего статистику времени перевода реализаций БП между состояниями, текущего и общего числа реализаций БП в различных состояниях (нагрузка на сотрудников, на БП). Варианты отклика планируется дополнить индикаторами сети. На рис. 2 представлены графики среднего времени выполнения и завершения реализации БП для случайного сотрудника и случайного БП. Выполнение соответствует непосредственной работе сотрудника над реализацией БП, завершение – существованию реализации от момента готовности к исполнению до окончания работы всех назначенных сотрудников.

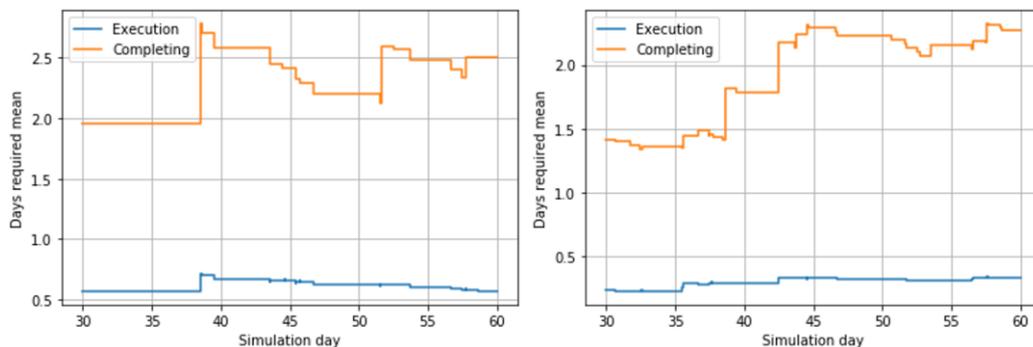


Рис. 2. Среднее время выполнения и завершения реализаций БП для случайного сотрудника (слева), случайного БП (справа)

Полученные отклики используются в модуле анализа чувствительности, выполняющем сопоставление неопределённости отклика модели к неопределённости во входных параметрах [4]. Ими, помимо параметров генерации БП и цепочек БП (определяющих поток заявок в систему), могут быть параметры исполнения БП отдельными сотрудниками, параметры, определяющие взаимодействие сотрудников, а также другие процессы, протекающие в компании.

На данном этапе завершена работа по реализации функционала мультиагентного имитационного моделирования, ввода данных, генерации и дополнения данных их синтетическими, отслеживания и анализа отклика. Выполнены численные эксперименты на дополненных реальных и полностью синтетических данных. Основные направления дальнейшей работы включают реализацию интерфейса для задания пользовательских слоёв МС и функций, используемых для моделирования взаимодействия сотрудников, а также, в качестве примера, провести исследование влияния процессов запроса, оказания помощи на выполнение БП и эволюцию МС.

### Литература

1. Law A.M. Simulation Modeling and Analysis. McGraw-Hill Education. 2015. 743 p.
2. Lieder M., Asif F.M.A., Rashid A. Towards Circular Economy implementation: an agent-based simulation approach for business model changes // Auton. Agent. Multi. Agent. Syst. 2017.
3. Battiston F., Nicosia V., Latora V. The new challenges of multiplex networks: Measures and models // Eur. Phys. J. Spec. Top. 2017.
4. Velumani S., Tang H. Operations Status and Bottleneck Analysis and Improvement of a Batch Process Manufacturing Line Using Discrete Event Simulation // Procedia Manuf. 2017.

**Ле Ань Ту**

Университет ИТМО,  
факультет инфокоммуникационных технологий,  
аспирант, 2-й год, группа 7725,  
направление подготовки: 05.12.14 – Радиолокация и радионавигация,  
e-mail: leanhtutcdt@gmail.com

УДК 004.82

## **СРАВНИТЕЛЬНЫЙ МЕТОД ОБНАРУЖЕНИЯ ДВИЖУЩИХСЯ НАЗЕМНЫХ ЦЕЛЕЙ**

**Ле Ань Ту**

### **Аннотация**

В настоящее время процесс обнаружения и визуализации нескольких движущихся целей с использованием радиолокационной синтезированной апертуры (SAR) становится все более актуальным. Наша задача заключается в анализе методов обнаружения движущихся наземных целей SAR, таких как индикация наземных движущихся целей (GMTI), визуализация наземных движущихся целей (GMTIm), метод обнаружения теней, нейронная сеть и так далее. В статье рассмотрены некоторые современные методы, благодаря которым возникает возможность улучшения эффективного обнаружения движущихся наземных целей.

### **Ключевые слова**

Индикация наземных движущихся целей (GMTI), визуализация наземных движущихся целей (GMTIm), метод обнаружения теней, искусственные нейронные сети, SAR.

В настоящее время задача обнаружения и визуализации нескольких движущихся целей, например, танков или колесных транспортных средств, с использованием радиолокационной синтезированной апертуры (SAR) становится все более актуальным. Точное обнаружение имеет решающее значение для измерения скорости цели, оценки местоположения и получения сфокусированных изображений. Поэтому для корректного определения SAR наземных объектов в условиях сильных помех требуется разработка методики обнаружения движущихся целей.

Целью исследования является анализ и выбор подходящих методов обнаружения движущихся наземных целей SAR.

На сегодняшний день уже существуют методы для обнаружения движущихся наземных целей SAR, такие как индикация наземных движущихся целей (GMTI), визуализация наземных движущихся целей (GMTIm), метод обнаружения теней, нейронная сеть и так далее. Рассмотрим каждый из аналогов поподробнее.

В статье [1] описан процесс обнаружения движущихся наземных целей SAR и предлагается новая совместная метрика амплитуды и фазы интерферограммы (IMP), где метрика IMP представляет из себя разницу в обратном рассеянии между движущимися целями и неподвижными помехами, содержащиеся в интерферограмме.

Путем анализа разницы в обратном рассеянии между движущимися целями и неподвижными помехами, содержащими в интерферограмме, в данной статье был предложен метод обнаружения постоянной частоты ложных тревог (CFAR), основанный на совместной метрике IMP SAR для наземных движущихся целей. Во-первых, используя исключительную взаимосвязь отображения между интерферометрической фазой и модулем соответствующего вектора и комбинируя интерферометрическую величину, была построена новая совместная метрика, называемая метрикой IMP. Во-вторых, в рамках мультипликативной модели,

основанной на комплексном распределении Wishart, была получена статистическая модель показателя IMP, которая обозначается как распределение  $S_0$ . Между тем, параметрические оценки распределения  $S_0$  также были представлены на основе преобразования Mellin. Наконец, аналитически задан порог CFAR для обнаружения наземных движущихся целей с использованием распределения  $S_0$  (рис. 1).

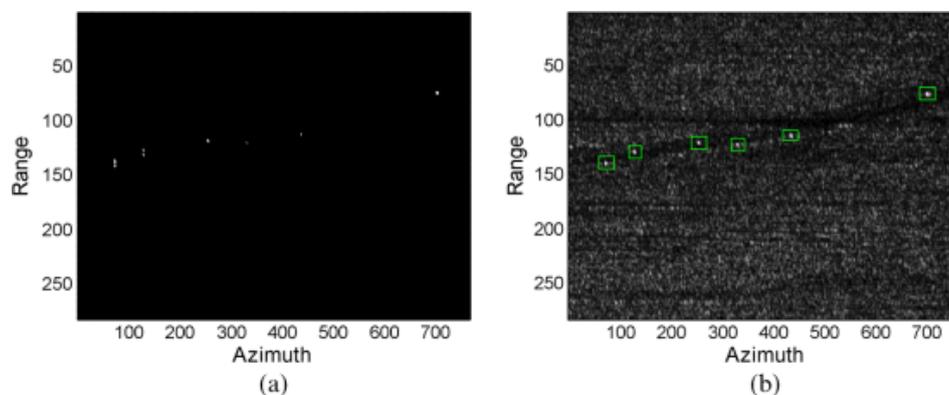


Рис. 1. Результаты обнаружения; а) Изображение результата обнаружения в двоичном формате, б) Результаты обнаружения соответствовали исходному изображению SAR

Эксперименты, проведенные с измеренными изображениями с двойным SAR, не только показывают эффективность статистических моделей метрики IMP и параметрических оценок, но также доказывают хорошие характеристики предлагаемого метода обнаружения CFAR.

В работе [2] движущиеся цели с различными параметрами движения классифицируются на основе относительного расположения их спектров по отношению к спектру помех. На основе классификации предлагается новая стратегия обработки движущихся целей с реальными данными SAR, включая двухэтапный алгоритм индикации движущихся по земле целей (GMTI) и практический алгоритм построения визуализации наземных движущихся целей (GMTIm) с компенсацией ошибок движения. Двухэтапный алгоритм GMTI может указывать на несколько движущихся целей, в частности на те, которые находятся под завалами. Практический алгоритм GMTIm показывает высокую производительность при построении изображения движущейся цели в реальных данных, поскольку ошибки движения оцениваются и компенсируются.

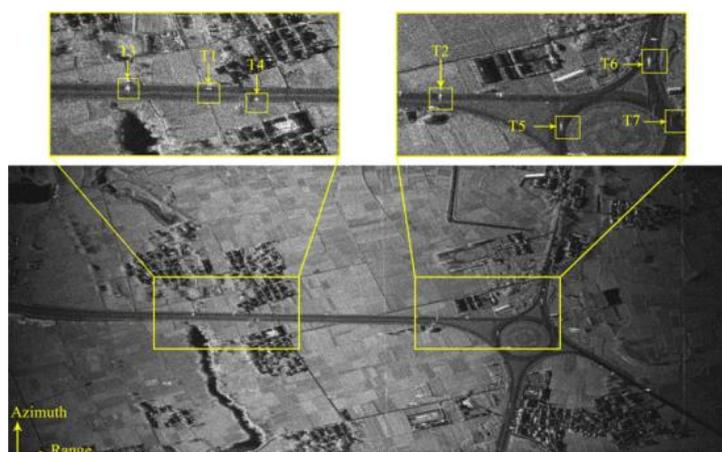


Рис. 2. Отображение движущихся целей на стационарном изображении с помощью предложенной стратегии [2]

Окончательный результат визуализации 7 движущихся целей с предложенной стратегией обработки движущихся целей показан на рис 2. Все движущиеся цели (отмеченные стрелками и прямоугольниками) обнаруживаются предложенной стратегией, а также они фокусируются и перемещаются в свои фактические локации. Таким образом, эффективность предложенной стратегии в реальных условиях доказана.

В исследовании [3] предложили эффективный метод обнаружения теней, основанный на слиянии нескольких элементов, который улучшает производительность обнаружения теней, и разработали модифицированную стратегию для дальнейшего повышения производительности обнаружения метода с использованием теней.

При увеличении дальности наблюдения радара многоканальная радиолокационная система высокого разрешения с синтезированной апертурой может страдать от снижения отношения сигнал / шум (SNR) цели, что приводит к ухудшению характеристик обнаружения наземной GMTI [4]. Функция тени, помимо амплитуды и интерферометрической фазы движущейся цели, может быть доступна для улучшения характеристик обнаружения цели.

Также очевидно, что область тени 3 настолько мала, что ее трудно обнаружить. Напротив, можно предсказать, что теневые области 1 и 2 имеют более высокую вероятность обнаружения из-за их большей занятости.

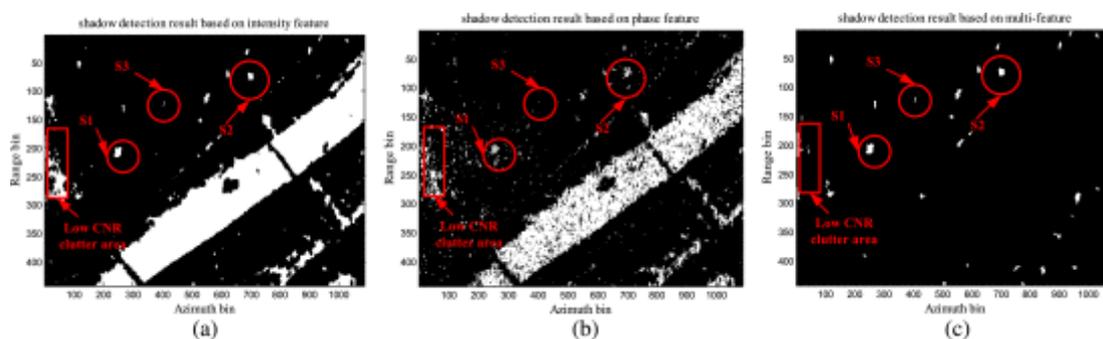


Рис. 3. Результаты обнаружения теней основаны на (а) только характеристике интенсивности изображения SAR, (б) только характеристике интерферометрической фазы SAR и (с) множественных характеристиках, соответственно

Результат обнаружения тени методом Оцу проиллюстрирован на рис. 3 (а), который основан на нормированной многопозиционной интерферометрической амплитуде. Очевидно, что количество областей ложной тени уменьшается, если объединить априорный размер и интерферометрические фазовые характеристики, как показано на рис. 3 (с). Кроме того, область помех с низким SNR, показанная на рис. 3 (а), очевидно, устраняется с помощью метода, основанного на слиянии множества элементов.

Наконец, результаты численного моделирования показывают, что теневой метод имеет лучшие характеристики обнаружения по сравнению с традиционными алгоритмами обнаружения. Однако также показано, что предложенный способ с использованием теней страдает от ухудшения характеристик обнаружения в области с высоким SNR, когда вероятность обнаружения для теневой области, особенно для теневой области с небольшим размером. Это ограничивает применение метода.

Автор [5] описал метод, позволяющий определять тип летательного аппарата в автоматическом режиме при обработке оптико-электронных разведывательных изображений. Метод основан на нейронной сети технологии распознавания изображений с обработкой заранее полученных данных самой сетью. Это дает возможность увеличить вероятность правильного распознавания, а также учесть возможные ошибки точных параметров.

Применение метода позволит совместить формализованную специфику воздушных объектов, которая может быть определена на основе полученных в автоматическом режиме изображений, и общую форму воздушного объекта. С помощью нейронной сети можно создавать поля, позволяющие не учитывать ошибки измерения (рис. 4).



Рис. 4. Важность функций

Эксперимент обнаружил, что только в трех случаях из ста объект на изображении не распознавался правильно, а в двух случаях из сотен объект распознавался при его отсутствии на изображении. В результате использования предложенного метода стало возможным получение стабильных результатов. Экспериментального набора достаточно, чтобы подтвердить вероятность правильного распознавания (рис. 5).

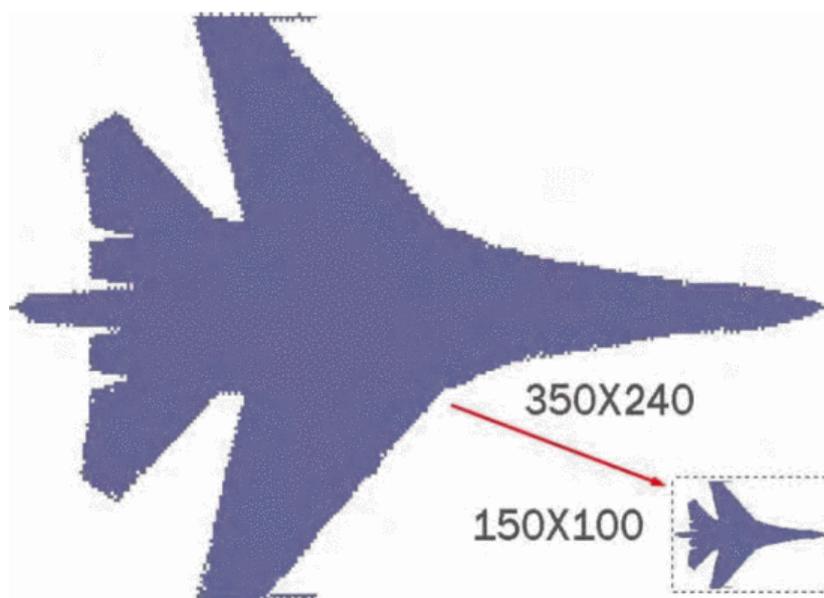


Рис. 5. Масштабирование выделенного воздушного объекта

Этот метод является оптимальным для решения задачи отнесения объекта к определенному классу. Реализация подобного метода позволит обрабатывать изображения в автоматическом режиме прямо на борту беспилотного летательного

аппарата или самолета и доставлять на наблюдательную станцию не снимки, а качественное и количественное соединение объектов на аэродроме.

В результате выполненной работы рассмотрены методы обнаружения движущихся наземных целей SAR. На основе полученных результатов определен оптимальный метод обнаружения движущихся наземных целей SAR. Разработанный метод работает на основе нейронных сетей и показывает высокую точность обнаружения наземных целей SAR. Таким образом, метод пригоден и может широко применяться в будущем для обнаружения движущихся наземных целей SAR.

### Литература

1. Gao G. and Shi G., "The CFAR Detection of Ground Moving Targets Based on a Joint Metric of SAR Interferogram's Magnitude and Phase," in *IEEE Transactions on Geoscience and Remote Sensing*, vol. 50, no. 9, pp. 3618-3624, Sept. 2012, doi: 10.1109/TGRS.2012.2184836.
2. Yang J., Liu C. and Wang Y., "Detection and Imaging of Ground Moving Targets With Real SAR Data," in *IEEE Transactions on Geoscience and Remote Sensing*, vol. 53, no. 2, pp. 920-932, Feb. 2015, doi: 10.1109/TGRS.2014.2330456.
3. Xu H., Yang Z., Chen G., Liao G. and Tian M., "A Ground Moving Target Detection Approach Based on Shadow Feature With Multichannel High-Resolution Synthetic Aperture Radar," in *IEEE Geoscience and Remote Sensing Letters*, vol. 13, no. 10, pp. 1572-1576, Oct. 2016, doi: 10.1109/LGRS.2016.2597159.
4. Sun H., Shimada M. and Xu F., "Recent Advances in Synthetic Aperture Radar Remote Sensing-Systems, Data Processing, and Applications," in *IEEE Geoscience and Remote Sensing Letters*, vol. 14, no. 11, pp. 2013-2016, Nov. 2017, doi: 10.1109/LGRS.2017.2747602.
5. Ruban I., Smelyakov K., Vitalii M., Dmitry P. and Bolohova N., "Method of neural network recognition of ground-based air objects," 2018 IEEE 9th International Conference on Dependable Systems, Services and Technologies (DESSERT), Kiev, 2018, pp. 589-592, doi: 10.1109/DESSERT.2018.8409200.



**Липатова Дарья Сергеевна**  
Университет ИТМО,  
Институт дизайна и урбанистики,  
студентка группы № С42042,  
направление подготовки: 09.04.03 – Прикладная информатика,  
e-mail: lipatovads@gmail.com

УДК 316.422

## **ВНЕДРЕНИЕ ЭЛЕКТРОННЫХ ГОСУДАРСТВЕННЫХ СОЦИАЛЬНЫХ УСЛУГ В ПРОАКТИВНОМ ФОРМАТЕ В САНКТ-ПЕТЕРБУРГЕ**

**Д.С. Липатова**

**Научный руководитель – зам. директора центра технологий  
электронного правительства Д.Р. Трутнев**

Работа выполнена в рамках темы НИР «Анализ данных о получателях государственных социальных услуг для первоочередного перевода в проактивную форму в Санкт-Петербурге».

### **Аннотация**

В работе представлен анализ зарубежного и регионального опыта проактивных государственных услуг, в процессе которого были выявлены причины и ожидаемые эффекты внедрения проактивных государственных услуг в Санкт-Петербурге, а также разработаны рекомендации по переводу электронных государственных социальных услуг в проактивную форму в Санкт-Петербурге.

### **Ключевые слова**

Государственные услуги, социальные услуги, проактивные услуги.

Портал «Государственные и муниципальные услуги Санкт-Петербурга» – единый портал, где расположена информация о порядке получения государственных и муниципальных услуг и где есть возможность подачи заявлений на получение этих услуг в онлайн-режиме на территории Санкт-Петербурга [5].

Понятие «проактивные государственные услуги» появилась относительно недавно, но на данный момент его все чаще начинают обсуждать и внедрять в повседневную жизнь общества. Проактивные государственные услуги – услуги, которые предоставляются по факту получения гражданином права на них без заявлений с его стороны, на основании сведений, полученных из государственных реестров и информационных систем.

Актуальность исследования основывается на федеральном проекте «Цифровое государственное управление» национальной программы «Цифровая экономика Российской Федерации». В данной программе одной из задач является разработка и принятие норм предоставления муниципальных услуг в проактивном формате, которые будут закрепляться в нормативно правовых актах [4].

Был проведен анализ зарубежного и регионального опыта внедрения проактивных государственных услуг. В 2019 году в Эстонии была внедрена проактивная социальная услуга, основанная на том, что нужные для ее получения данные уже есть в органах власти и получателям не придется предоставлять их повторно (например, семейные

пособия по жизненной ситуации «Рождение ребенка»). В соответствии с этим принципом в Эстонии предоставляются услуги, связанные с семейными пособиями [1]. Также в 2019 году в Кировской области заработала ежемесячная региональная услуга по выплате 5000 рублей на второго ребенка [2].

С 15 апреля 2020 года Пенсионный фонд приступил к проактивному выдачу сертификатов на материнский капитал. Сведения о ребенке появляются в ЕГР ЗАГС, далее данные поступают в Пенсионный фонд. Сотрудники Пенсионного фонда проверяют наличие возможных ограничений, и, если проблем не найдено, происходит оформление сертификата. Получателю госуслуги в личном кабинете на сайте Пенсионного фонда или Едином портале государственных услуг приходит уведомление о готовом документе. На сегодняшний момент в Санкт-Петербурге и Ленинградской области в проактивном режиме оформлено свыше 18500 сертификатов на материнский капитал [6].

Для проактивного назначения ежемесячной денежной выплаты инвалидам и детям-инвалидам специально был создан Федеральный реестр инвалидов. Ежемесячная денежная выплата устанавливается со дня признания человека инвалидом или ребенком-инвалидом, и информация передается сотрудниками бюро медико-социальной экспертизы в Федеральный реестр инвалидов. Далее сведения в Пенсионный фонд поступают из Федерального реестра инвалидов, и происходит назначение выплаты. От получателя требуется только заявление о предпочитаемом способе доставке ежемесячной денежной выплаты, которое подается через личный кабинет на Едином портале государственных услуг или сайте Пенсионного фонда. Ранее услуга требовала личной подачи документов в МФЦ [3].

Благодаря анализу регионального и зарубежного опыта были выявлены причины внедрения проактивных электронных государственных услуг в Санкт-Петербурге:

- необходимость снижения бумажного документооборота;
- необходимость сокращения времени предоставления государственных услуг;
- необходимость решения проблемы кадрового дефицита.

Также были выявлены ожидаемые эффекты от внедрения услуг в проактивном формате в Санкт-Петербурге:

- сокращение времени получения государственных услуг;
- получение государственных услуг без заявлений получателя услуг;
- рост количества оказанных услуг;
- сокращение количества бумажных документов;
- удовлетворенность граждан.

На портале государственных услуг Санкт-Петербурга сведения о получателе услуг доступны из следующих ведомственных систем:

1. МВД России по городу Санкт-Петербургу и Ленинградской области (пол, сведения о паспорте, гражданстве, водительском удостоверении, регистрации гражданина по месту жительства или месту пребывания).
2. Управление Федеральной налоговой службы России по Санкт-Петербургу (идентификационный номер налогоплательщика, сведения о доходах и суммах налога физического лица и др.).
3. Межрайонная инспекция Федеральной налоговой службы России по Санкт-Петербургу.
4. Территориальное управление ПФР в Санкт-Петербурге (СНИЛС, сведения о назначенных пенсиях и социальных выплатах, состоянии индивидуального страхового счета застрахованного лица, сведения из электронной трудовой книжки).
5. Территориальный фонд ОМС Санкт-Петербурга (сведения о Едином номере полиса обязательного страхования и страховой принадлежности застрахованного лица).

В базы территориальных органов ведомственных систем Санкт-Петербурга сведения о гражданине поступают через ведомственные системы РФ. Вышеуказанные сведения доступны через портал при наличии таких сведений у гражданина (например, наличие транспортных средств, трудовой книжки, мер социальной поддержки и т. д.) и с его разрешения на запрос.

Был проведен анализ административных регламентов предоставления государственных услуг в Санкт-Петербурге, в результате которого был выделен перечень документов и условий, необходимых для получения каждой услуги. В том числе были проанализированы результаты и документы, которые гражданин получит в ходе обработки госуслуг. На основе данного анализа были выделены триггеры, которые могут способствовать получению государственных услуг гражданином, и после выделен перечень уникальных триггеров (табл. 1).

Таблица 1

### Уникальные триггеры

| Триггеры  | Способ получения                               |
|---|--|
| Возраст   | Данные из ЕСИА                                 |
| Получение документа, подтверждающий, принадлежность к категории детей сирот и детей, оставшихся без попечения родителей       | Результат раннее оказанной услуги (ОСЗН, ПФР)  |
| Получение сертификата материнского капитала   | Результат раннее оказанной услуги (ПФР)        |
| Получение медицинского заключения (заболевания и пр.)   | Указываются заявителем                         |
| Необходимость в покупке/ремонте жилого помещения  | Указывается заявителем                         |
| Получение пенсионного удостоверения   | Результат раннее оказанной услуги (ПФР)        |
| Поступление ребенка в ОУ  | Результат раннее оказанной услуги (ОУ)         |
| Получение документа, которое подтверждает статус опекуна или попечителя, опекуна или попечителя, исполняющие свои обязанности | Результат раннее оказанной услуги (ЗАГС, ОСЗН) |
| Получение свидетельства об усыновлении  | Результат раннее оказанной услуги (ЗАГС)       |
| Получение свидетельства о смерти  | Результат раннее оказанной услуги (ЗАГС)       |
| Получение сертификата многодетной семьи   | Результат раннее оказанной услуги (МФИЦ)       |
| Получение удостоверения инвалида  | Результат раннее оказанной услуги (ОСЗН)       |
| Получение статуса малообеспеченной семьи  | Результат раннее оказанной услуги              |
| Получение свидетельства о рождении ребенка  | Результат раннее оказанной услуги (ЗАГС)       |
| Получение статуса беременной женщины  | Указывается заявителем                         |

Большинство уникальных триггеров возникают на основе услуг и их результатов, которые предоставлялись ранее заявителю на портале государственных услуг Санкт-Петербурга.

В результате проведенного исследования для внедрения проактивного формата предоставления электронных государственных социальных услуг в Санкт-Петербурге рекомендуются следующие действия:

1. Внести изменения в Федеральный закон от 14 июля 2010 года № 210-ФЗ «Об организации предоставления государственных и муниципальных услуг» статья 4 пункт 2 (добавить беззаявительный порядок обращения за предоставлением государственных и муниципальных услуг).

2. Создать единую базу малоимущих семей (по примеру работы ФРИ).
3. Создать единую базу по составу семьи (по примеру работы ФРИ).
4. Автоматизировать межведомственные запросы.

В созданной базе малоимущих семей и базе по составу семьи должно быть постоянное обновление информации о получателях услугах, в том числе и проверка возможного продления услуг у гражданина, если он ранее уже получал услугу (например, для инвалидов, малоимущих и т. д.). Единые базы будут хранить информацию о гражданах, тем самым гражданину при получении услуг не потребуются предоставлять необходимые документы самостоятельно. Сотрудникам государственных и муниципальных организаций не придется самостоятельно готовить запрос и ответ на межведомственные запросы. Полноценное автоматизированное межведомственное взаимодействие повлияет на сокращение срока оказания услуги заявителю, тем самым будет соответствовать выявленным ранее эффектам внедрения электронных проактивных услуг в Санкт-Петербурге.

Рекомендованные изменения повлияют на сокращение срока оказания услуги заявителю (отсутствие ручного межведомственного запроса, наличие полноценного межведомственного взаимодействия органов государственной власти), тем самым приведет к ожидаемым эффектам внедрения электронных проактивных услуг в Санкт-Петербурге.

Благодаря анализу нормативных документов, данных о получателях государственных услуг в Санкт-Петербурге и проактивного подхода предоставления услуг были составлены рекомендации по внедрению электронных государственных социальных услуг в проактивном формате в Санкт-Петербурге и электронные государственные услуги, рекомендованные для перевода в проактивный формат в Санкт-Петербурге.

Исследовательская работа выполнялась совместно со СПб ГУП «Санкт-Петербургским информационно-аналитическим центром».

## Литература

1. Департамент социального страхования Эстонии. Виды семейных компенсаций [Электронный ресурс]. Режим доступа: <https://www.sotsiaalkindlustusamet.ee/ru/deti-semi/vidy-semeunyh-kompensaciy> (дата обращения: 10.01.2021).
2. Кировское областное государственное автономное учреждение МФЦ. Кировский проект «проактивных услуг» представили на конференции цифрового будущего в Москве [Электронный ресурс]. Режим доступа: <https://xn--43-jlcennldkесбсј0j.xn--plai/mfc/news/kirovskiy-proekt-proaktivnykh-uslug-predstavili-na-konferentsii-tsifrovogo-budushchego-v-moskve.html> (дата обращения: 10.01.2021).
3. Министерство труда и социальной защиты РФ. Родители 13 тыс. новорожденных получили номер СНИЛС ребенка без заявлений [Электронный ресурс]. Режим доступа: <https://mintrud.gov.ru/social/social/1226> (дата обращения: 10.01.2021).
4. Министерство цифрового развития, связи и массовых коммуникаций РФ. Федеральный проект «Цифровое государственное управление» национальной программы «Цифровая экономика РФ» [Электронный ресурс]. Режим доступа: <https://digital.gov.ru/ru/activity/directions/858/> (дата обращения: 10.01.2021).
5. Портал государственных услуг Санкт-Петербурга [Электронный ресурс]. Режим доступа: <https://gu.spb.ru/> (дата обращения: 10.01.2021).
6. Портал Муниципального округа Васильевский. Что такое проактивное оформление материнского капитала? [Электронный ресурс]. Режим доступа: <http://msmov.spb.ru/жителям-округа/smi-municipalnogo-okruga/informacionnaya-lenta/pfr168.html> (дата обращения: 10.01.2021).



**Логунов Алексей Андреевич**

Год рождения: 1997

Университет ИТМО,

факультет информационных технологий и программирования,  
студент группы № М41211,

направление подготовки: 09.06.01 – Информатика

и вычислительная техника,

e-mail: khabalexlogunov@gmail.com



**Новоселов Сергей Александрович**

ЦРТ-Инновации,

научно-исследовательский департамент,

в.н.с.

УДК 004.934

## **СРАВНЕНИЕ МЕТОДОВ ПОТОКОВОЙ ДИАРИЗАЦИИ**

**А.А. Логунов**

**Научный руководитель – в.н.с. С.А. Новосёлов**

Работа выполнена в рамках темы НИР №620172 «Определение структуры полилога с применением лексических и нелексических признаков речи нескольких дикторов».

### **Аннотация**

В работе рассмотрены алгоритмы потоковой диаризации дикторов за счет применения рекуррентных нейронных сетей и моделей случайных процессов. Проведены замеры и сравнение качества представленных моделей с результатами работы модели офлайн диаризации.

### **Ключевые слова**

Нейронные сети, обработка речи, диаризация.

Процесс обработки речевого сигнала может быть направлен на решение большого количества различных задач, одной из которых является диаризация.

Диаризация – это процесс сегментации аудиоданных во временной области в соответствии с тем, какой говорящий в каждый момент времени говорит. Процесс диаризации не ставит во главу угла определение идентификатора говорящего. Задача состоит в том, чтобы поставить метку кластера каждому диктору в каждый момент времени. Таким образом, метка лишь определяет номер говорящего, не привязываясь к конкретным личностям дикторов.

Диаризацию обычно используют как один из этапов предобработки данных в конвейере с другими алгоритмами, скажем, для автоматического распознавания речи или её транскрибирования. Типичная задача с применением диаризационных алгоритмов – протоколирование встреч, заседаний или собраний.

Общий алгоритм работы конвейеров диаризации может быть описан последовательными шагами предобработки сырых аудио-данных, делением их на окна фиксированной длины, извлечением высокоуровневых векторных представлений, их

агрегированием в сегменты временной области и кластеризацией сегментов. Схема этого процесса показана на рисунке.

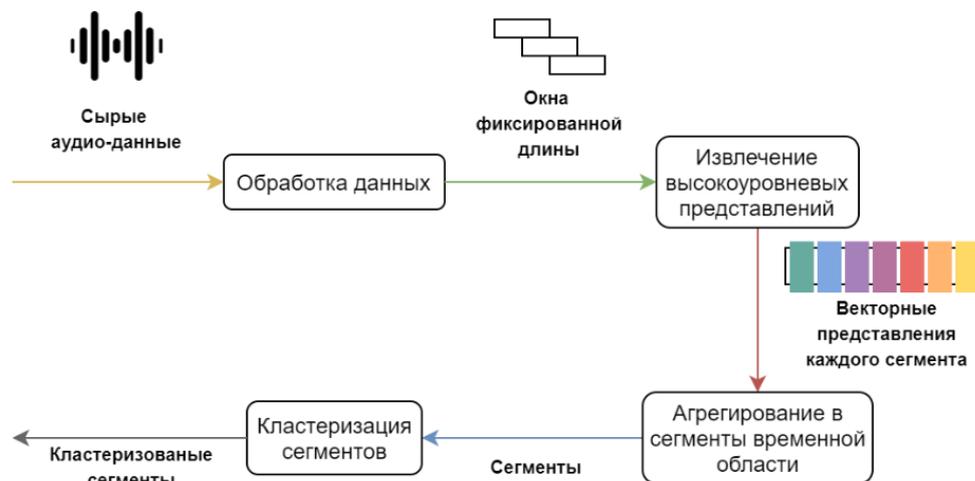


Рисунок. Схема конвейера стандартных алгоритмов диаризации

Для определения принадлежности сегментов речи к какому-либо из говорящих стандартные конвейеры диаризации используют алгоритмы машинного обучения без учителя. Например, алгоритмы спектральной или агломеративной кластеризации. Соответственно, задача диаризации начинает лежать в области обучения без учителя. В некоторых случаях существует необходимость применения таких систем к поступающему в режиме онлайн аудио потоку данных. Как пример можно привести диаризацию дикторов на стриминговых сервисах.

Очевидное применение уже имеющихся систем к решению такого вида задачи может привести к неудаче по одной из следующих причин:

1. Использование непотоковых алгоритмов кластеризации, которым нужно знать информацию обо всех дикторских векторах, участвующих в произнесении.
2. Будучи представителями категории алгоритмов машинного обучения без учителя, стандартные алгоритмы кластеризации не используют кластерные метки дикторов, доступные в данных.
3. Игнорирование алгоритмами кластеризации темпоральности данных – то есть вытянутости данных во времени.

В табл. 1. показаны возможные способы решения сложившихся проблем и вытекающих из них отрицательных последствий.

В исследовании от компании Google рассматривается подход к диаризации с использованием рекуррентной нейронной сети с неограниченными чередующимися состояниями (unbounded interleaved-state recurrent neural network, UIS-RNN) [1]. Неограниченность относится к тому, что система не ограничена заранее заданным количеством дикторов в произнесении и способна сама, в режиме онлайн, определять их число.

Ключевое отличие предложенного метода от рассмотренных выше состоит в том, что дикторские представления (embeddings) моделируются рекуррентной нейронной сетью с совместным использованием параметров, а разделение дикторов происходит с использованием различных состояний RNN, которые чередуются во временной области. То есть, уникальными для каждого диктора в рекуррентной модели являются скрытые состояния (hidden states), а общими – тренируемые параметры (веса, theta)

**Подходы к решению проблем**

| Проблема  | Решение  | Последствия   |
|---|--|---|
| Неиспользование кластерных меток в обучающих данных                       | Использование алгоритмов кластеризации с частичным привлечением учителя  | Не учитывается темпоральность данных  |
| Неиспользование темпоральной информации в данных                          | Использование моделей статистических процессов - например, скрытой марковской модели (hidden Markov model HMM) | Подобные модели используют строгие и достаточно наивные предположения. Модели хорошо работают, если предположения так же сбываются на поступающих на вход данных    |
| Невозможность донастройки параметров уже созданной модели в режиме онлайн | Использование онлайн-модификаций стандартных алгоритмов кластеризации  | Не учитывается темпоральность данных. Однако все-таки существуют онлайн-адаптации вроде потоковой скрытой Марковской модели (OnlineHMM, online hidden Markov model) |

Использование же меток дикторов позволяет обучить предложенную модель, используя стандартные алгоритмы стохастического градиентного спуска. В результате, модель может быть использована для диаризации на новых произнесениях от неизвестных ранее дикторов. Наличие подобного свойства у предложенной системы может быть очень полезно в отношении определенных редких языков, для которых нет достаточного количества тренировочных данных.

Общий алгоритм работы этой системы может быть разбит на следующие шаги:

1. Определение, является ли текущий диктор незнакомым для системы. Используется модель процесса китайского ресторана (chinese restaurant process, CRP). Это, в частности, позволяет применить алгоритм к потоковым данным с сохранением должного качества.

2. Создание своего экземпляра модели со своим начальным состоянием  $h_0$  для каждого нового диктора.

3. Обновление скрытого состояния модели текущего диктора происходит до тех пор, пока этот диктор не прекратит говорить.

4. При переключении на другого диктора в другую рекуррентную модель переносится состояние о весах прежней модели, но не ее скрытое состояние.

Такой алгоритм использует кластерные метки в данных при обучении, учитывает темпоральную структуру данных за счет использования рекуррентной нейронной модели в своей структуре и помогает решить задачу потоковой диаризации.

Тем не менее, качество данной системы удалось улучшить путем применения функции потерь выборочного среднего и улучшения модели китайского ресторана за счет применения некоторых эвристик. В конечном счете, полученная модель базируется на исходной. Полученная система получила название рекуррентной нейронной сети с неограниченными чередующимися состояниями и функцией потерь выборочного среднего (unbounded interleaved-state recurrent neural network with sample mean loss, UIS-RNN-SML) [2-4].

Примененные к CRP эвристики касаются свойств человеческой речи и человеческого диалога и настройки статистических параметров модели CRP.

Обе описанные модели были протестированы на тестовой части набора данных DIHARD II с подсчетом метрики ошибки диаризации (diarization error rate, DER). Причем подсчет происходил для двух вариантов данных: с наложениями и без них. Под наложениями понимаются части полилога, в рамках которых одновременно говорит несколько дикторов. Типичным примером наложений являются перебивания одним диктором другого.

Для сравнения также представлены результаты работы алгоритма вероятностного линейного дискриминантного анализа с добавлением агломеративной иерархической кластеризации (probabilistic linear discriminant analysis with agglomerative hierarchical clustering, PLDA+АНС) как типичного представителя оффлайн-кластеризации.

Результаты вычисления метрики представлены в табл. 2.

Таблица 2

### Сравнение качества моделей

| Метод       | DER  | DER (без наложений) |
|-------------|------|---------------------|
| UIS-RNN     | 30.9 | 23.4                |
| UIS-RNN-SML | 27.3 | 19.4                |
| PLDA+АНС    | 26.1 | 17.7                |

Из табл. 2. очевидно следует то, что методы онлайн диаризации проигрывают в качестве аналогу на базе агломеративной кластеризации. Кроме того, улучшение алгоритма UIS-RNN с SML почти позволяет добиться качества оффлайн-алгоритма, показывая отставание по метрике DER примерно в 1.5%.

В ходе работы были рассмотрены методы диаризации на основе рекуррентных нейронных сетей, способные вовлечь в тренировку кластерные метки, доступные в данных и принимающие во внимание их темпоральную структуру. На рассмотренных алгоритмах был проведен расчет качества по метрике DER. Из результатов следует, что онлайн-методы диаризации уступают в качестве оффлайн-алгоритмам на небольшой процент по метрике.

### Литература

1. Fully Supervised Speaker Diarization [Электронный ресурс]. Режим доступа: <https://arxiv.org/pdf/1911.01266.pdf> (дата обращения: 23.02.2021).
2. Supervised online diarization with sample mean loss for multi-domain data [Электронный ресурс]. Режим доступа: <https://arxiv.org/pdf/1911.01266.pdf> (дата обращения: 23.02.2021).
3. The Second DIHARD Diarization Challenge: Dataset, task, and baselines [Электронный ресурс]. Режим доступа: <https://arxiv.org/pdf/1906.07839.pdf> (дата обращения: 23.02.2021).
4. Accurate Online Speaker Diarization with Supervised Learning [Электронный ресурс]. Режим доступа: <https://ai.googleblog.com/2018/11/accurate-online-speaker-diarization.html> (дата обращения: 23.02.2021).



**Мамаев Никита Константинович**  
Университет ИТМО,  
факультет информационных технологий  
и программирования,  
аспирант, группа 7717,  
Направление подготовки: 09.06.01 – Информатика и  
вычислительная техника,  
e-mail: nikita.mamaev1@gmail.com



**Матвеев Юрий Николаевич**  
Университет ИТМО,  
факультет информационных технологий и программирования,  
д. т. н., профессор,  
руководитель международной лаборатории «Многомодальные  
биометрические и речевые системы»,  
заведующий корпоративной лабораторией технологий  
человеко-машинного взаимодействия,  
e-mail: matveev@mail.ifmo.ru

УДК 004.89

## **ОЦЕНКА НЕУВЕРЕННОСТИ НЕЙРОННЫХ МОДЕЛЕЙ ДЛЯ ТЕКСТОВОЙ КЛАССИФИКАЦИИ**

**Н.К. Мамаев**

**Научный руководитель – д.т.н., профессор Ю.Н. Матвеев**

Работа выполнена в рамках темы НИР №620183 «Разработка виртуального диалогового помощника для поддержки проведения дистанционного экзамена на основе моделей-трансформаторов и понимания естественного и математического языка».

### **Аннотация**

Одной из наиболее проблем в сфере машинного обучения является проблема плохой калибровки значений уверенности в предсказании, что затрудняет определение недостоверных предсказаний и выбросов и, как следствие, делает модель менее надёжной. Цели текущей работы – привести некоторые способы решения этой проблемы, применить их к модели текстовой классификации и изложить результаты.

### **Ключевые слова**

Обработка естественного языка, текстовая классификация, определение выбросов, confidence.

Текстовая классификация – классическая задача направления автоматической обработки естественного языка (NLP). Алгоритмы, выполняющие текстовую классификацию, широко применяются при обработке текстовых и голосовых запросов в каналах поддержки (тематическая и интент-классификация), при изучении публичного мнения (анализ тональности), обработке публицистического контента и т.д. За последние два десятилетия методология текстовой классификации претерпела переход от парсеров и систем на правилах к методам машинного обучения, а затем – к глубоким нейронным сетям.

Для машинного обучения и НС характерна проблема плохой калибровки значений уверенности в предсказании. Это ведёт к двум последствиям: во-первых, сложно выделить предсказания, в которых модель «не уверена», то есть те, которые с высокой

вероятностью окажутся неверными. Во-вторых, при решении задачи открытой (open set) классификации тяжело идентифицировать случаи, когда входные данные не соответствуют распределению обучающих данных (случай, когда подаваемый на модель пример не является экземпляром ни одного из целевых классов). Последняя проблема – называемая обнаружением выбросов или out-of-distribution detection – была недавно признана значительным препятствием к построению надёжного искусственного интеллекта [1]. Она актуальна, например, при обучении тематического классификатора запросов клиента: необходимо отличить целевые реплики от так называемого оффтопа, под которым могут подразумеваться всяческие приветствия, запросы на незнакомую тему и с целью поболтать, а также запросы на иностранном языке. Решению этой проблемы и посвящена данная работа: основной задачей является классификация тематик запросов, а в качестве дополнительных «незнакомых» датасетов используются запросы по незнакомым темам, отзывы и новостные заголовки. В качестве нейронной модели используется свёрточная модель со слоем словных векторных представлений, для инициализации которого используются подготовленные векторы.

Одним из возможных путей к получению достоверных оценок уверенности – обучить модель выводить её явно и отказаться от классической оценки уверенности, т. е. наибольшего softmax-значения на выходе модели. Этот подход описывается в работе [2], и решение достигается за счёт встраивания дополнительного выходного слоя для оценки уверенности модели, параллельно с классификационным softmax-слоем. Модель также имеет возможность «воспользоваться подсказкой» – для этого вектор предсказаний смешивается с ground-truth вектором, а в качестве веса берётся новая оценка уверенности. За счёт нового слагаемого в функции потерь оценка уверенности устремляется к максимуму – таким образом, модель мотивирована пользоваться подсказкой только в ситуации неопределённости. В статье [2] этот подход успешно применяется к классификации изображений и демонстрирует естественное убывание уверенности по мере увеличения степени размытия, зашумления и добавления артефактов сжатия.

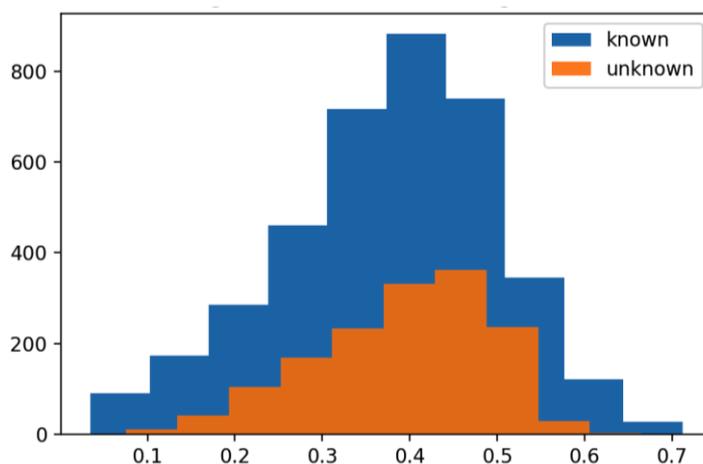


Рис. 1. Гистограмма значений уверенности при применении подхода, изложенного в [2] (синим – уверенность для целевых примеров, оранжевым – для нецелевых)

Результаты применения метода, изложенного в указанной работе, можно видеть на рис. 1. Гистограммы совпадают по форме (отличаются на множитель за счёт разного количества примеров в датасетах), что говорит о неразличимости целевых и нецелевых примеров.

Ещё одним подходом к генерации достоверных оценок уверенности является т.н. техника *mixup*: авторы работы [3] полагают, что она позволяет подавить алеаторическую неуверенность (*aleatoric uncertainty*), вызванную наличием шума во входных данных. Под *mixup*-обучением поднимается обучение модели на случайно взвешенных суммах пар примеров из одного минибатча. В качестве оценки уверенности в работе предлагается использовать расстояние Махаланобиса до распределения признаков каждого класса. Расстояние Махаланобиса традиционно применяется для определения выбросов, и для его вычисления необходимо посчитать своеобразные эталоны – средние признаков обучающих примеров, входящих в каждый класс, известный модели. В статье [3] этот подход применялся к такой же проблеме – текстовой классификации, но на английском языке.

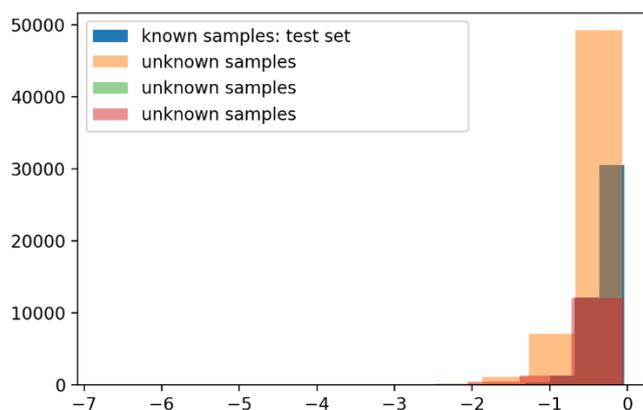


Рис. 2. Гистограмма значений уверенности при применении подхода, изложенного в [3] (синим – уверенность для целевых примеров, другими цветами – для нецелевых)

Как видно из рис. 2, распределение значений уверенности для целевых примеров остались практически неотличимым от распределения для нецелевых примеров, то есть применение этого подхода также не увенчалось успехом.

Наконец, достичь достоверных оценок уверенности можно также с помощью моделирования плотности. Подход, изложенный в работе [4], предполагает построение модели плотности вероятности распределения признаков для каждого класса; например, гауссовой смеси (GMM). С помощью модели гауссовой смеси оценивается правдоподобие того, что пример был порождён одним из поклассовых распределений (рис. 3).

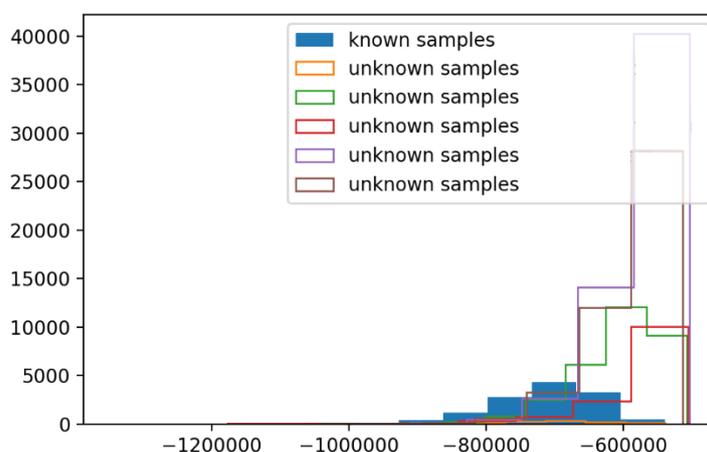


Рис. 3. Гистограмма значений уверенности при применении подхода, изложенного в [4] (синим – уверенность для целевых примеров, другими цветами – для нецелевых)

Для проведения экспериментов была использована однокомпонентная GMM. Этот метод позволил слегка отделить целевые примеры от нецелевых; тем не менее, достигнутого эффекта едва ли достаточно.

Таким образом, изложенные подходы к оценке уверенности модели не сработали на текущей задаче – с помощью новых оценок невозможно отделить целевые реплики от нецелевых. Причины этого результата могут быть следующие:

1. Использованные нецелевые (out-of-distribution) датасеты слишком похожи на целевой.

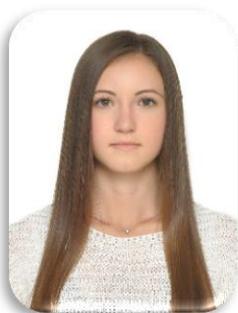
2. Предподготовленные словные векторные представления (эмбединги) мешают модели выделять нужные характеристики.

3. Неподходящая архитектура классификатора.

Перечисленные гипотезы будут проверены в рамках дальнейшей работы по данному исследованию.

### Литература

1. Amodei D., Olah C., Steinhardt J., Christiano P., Schulman J., Mané D. Concrete Problems in AI Safety. Режим обращения: <https://arxiv.org/abs/1606.06565> (дата обращения: 20.02.2021).
2. Terrance DeVries and Graham W. Taylor. Learning Confidence for Out-of-Distribution Detection in Neural Networks. Режим обращения: <https://arxiv.org/abs/1802.04865> (дата обращения: 20.02.2021).
3. Jianfeng He, Xuchao Zhang, Shuo Lei, Zhiqian Chen, Fanglan Chen, Abdulaziz Alhamadani, Bei Xiao, Chang-Tien Lu. Towards More Accurate Uncertainty Estimation In Text Classification. Режим обращения: <https://www.aclweb.org/anthology/2020.emnlp-main.671.pdf> (дата обращения: 20.02.2021).
4. Akshayvarun Subramanya, Suraj Srinivas, R.Venkatesh Babu. Confidence estimation in Deep Neural networks via density modelling. Режим обращения: <https://arxiv.org/abs/1707.07013> (дата обращения: 20.02.2021).



**Маркова Юлия Сергеевна**

Год рождения: 1998  
Университет ИТМО,  
факультет инфокоммуникационных технологий,  
студент группы № К4112с,  
направление подготовки: 11.04.02 – Инфокоммуникационные  
технологии и системы связи,  
e-mail: dragonfly\_jm@mail.ru



**Осипов Никита Алексеевич**

Год рождения: 1972  
Университет ИТМО,  
факультет инфокоммуникационных технологий,  
к.т.н., доцент,  
e-mail: nikita@ifmo.spb.ru

УДК 004.4

**АНАЛИЗ И ИССЛЕДОВАНИЕ БЫСТРОГО  
ВЫПОЛНЕНИЯ СЦЕНАРИЕВ ТЕСТИРОВАНИЯ  
ИНФОКОММУНИКАЦИОННОЙ СИСТЕМЫ**

**Ю.С. Маркова**

**Научный руководитель – к.т.н. Н.А. Осипов**

**Аннотация**

В работе рассмотрена оптимизация тестирования программного обеспечения и поиск возможных улучшений с целью экономии времени и затрат на данный процесс при совмещении ручных методов и автоматизации на некоторых этапах цикла тестирования разрабатываемой системы для максимальной эффективности процессов и обеспечения наибольшего качества продукта.

**Ключевые слова**

Тестирование программного обеспечения, оптимизация процессов тестирования, тестирование инфокоммуникационной системы, автоматизация, подходы к тестированию.

В современном мире любой выпускаемый программный продукт должен быть качественно протестирован. Однако подходов к тестированию программного обеспечения существует достаточно много. Не все они обеспечивают эффективный процесс тестирования, а также равномерное распределение ресурсов на данный процесс. Например, при использовании самого распространенного инструмента оптимизации тестирования, автоматизации выполняемых тестов, зачастую возникает проблема нехватки времени, денежных ресурсов, а также профессиональных кадров для полной реализации и поддержания процесса тестирования. В ходе работы рассматривается возможность улучшения описанного подхода к оптимизации процессов за счет совмещения ручного и автоматизированного тестирования. Данный способ позволит экономить время по сравнению с ручным прохождением сценариев, а также не тратить лишние ресурсы на полную автоматизацию всех выполняемых тестов.

Для того, чтобы выработать оптимальный подход к тестированию разрабатываемого программного обеспечения, необходимо рассмотреть основные существующие виды тестирования и применимость автоматизации процесса на основе задач, преследуемых в том или ином виде тестирования.

Например, как известно, Smoke-тестирование и регрессионное тестирование содержат в себе основные, базовые проверки функционала системы и предполагают частое прохождение одних и тех же сценариев. Это значит, что выполнение данного тестирования не оптимально производить вручную за счет повторяемости сценариев. Здесь Smoke-тестирование – это некоторый короткий цикл тестов для проверки самого базового функционала программной системы в каждой очередной выдаче сборки [1]. А регрессионное тестирование – это проверка существующего функционала на исправность, чтобы удостовериться, что новая часть программного кода не повлияла на уже разработанный функционал [2].

Еще одним примером служат нагрузочное и интеграционное тестирование. Нагрузочное тестирование представляет собой проверку системы на устойчивость в условиях большого количества входных данных [1]. А интеграционное - тестирование взаимодействия нескольких модулей системы или взаимодействия продукта со сторонними, внешними системами [2]. Это значит, что на создание тестовых данных, а затем и на выполнение самого сценария потребуется значительное количество времени, если проделывать данный процесс вручную.

По результатам анализа основных видов тестирования была составлена таблица, отражающая применимость и целесообразность ручного выполнения сценариев в том или ином виде тестирования.

Таблица

### Применимость и целесообразность ручного тестирования

| Виды тестирования             | Применимость ручного тестирования | Целесообразность ручного тестирования (по шкале от 0 до 10 субъективной оценки) |
|-------------------------------|-----------------------------------|---|
| Тестирование документации     | +                                 | 10  |
| Sanity-тестирование           | +                                 | 10  |
| Smoke-тестирование            | +                                 | 3   |
| Регрессионное тестирование    | +                                 | 4   |
| Интеграционное тестирование   | +                                 | 8   |
| Тестирование безопасности     | +                                 | 10  |
| Нагрузочное тестирование      | +                                 | 3   |
| Функциональное тестирование   | +                                 | 10  |
| Нефункциональное тестирование | +                                 | 9   |
| Usability-тестирование        | +                                 | 10  |

Из таблицы видно, что применение ручного тестирования возможно в любых видах тестирования, вопрос только в целесообразности – сколько времени и ресурсов уйдет на данный процесс. Для того, чтобы оптимально использовать временные ресурсы команды, необходимо совмещать два типа тестирования: использовать автоматизацию в тех случаях, когда это приведет к уменьшению рутинных задач, выполняемых вручную, и использовать ручное тестирование, когда требуется серьезная аналитическая работа или функционал подвержен большому количеству изменений [3].

Именно поэтому основной задачей работы является формулировка оптимальных условий тестирования для совмещения автоматизации и ручного метода.

Для того, чтобы сформулировать предложения по совершенствованию процесса тестирования, кратко рассмотрим пример инфокоммуникационной системы – программный продукт по продаже услуг связи абонентам. Предположим, разрабатывается система, позволяющая сформировать заказ пользователя для подключения желаемых услуг и отслеживать готовность данного заказа. На входе мы также можем иметь различные варианты тестовых данных, например, такие как тип клиента, тип заключаемого договора, подключаемые услуги и т.д.

Основной цикл тестирования представленной системы состоит из следующих этапов:

1. Подготовка тестовых данных, необходимых для выполнения сценария. (Test Data Creation).
2. Непосредственное прохождение сценария. (Execution).
3. Массовая проверка каких-либо полученных результатов. (Bulk Validation).
4. Анализ найденных проблем. (Triage & Troubleshooting).

Для того, чтобы инженер мог вручную проходить сценарии, но при этом не тратить время на лишние задачи внутри цикла тестирования, предлагается частично автоматизировать с помощью программных средств некоторые этапы.

### **Test Data Creation**

Подготовка тестовых данных для одного сценария может занять значительную часть всего времени прохождения. Особенно если начальными данными для одного сценария является весь пройденный цикл другого (законченный end-to-end процесс). В этом случае предлагается использовать автоматизацию подготовки данных для последующей работы с ними.

### **Bulk Validation**

При прохождении сценариев часто можно столкнуться с большим количеством данных, которые необходимо проверить на корректность в соответствии с документацией. Особенно часто такая необходимость возникает в интеграционном тестировании, когда нужно сравнить ожидаемый и реальный запрос и ответ. Для этого можно разработать и использовать программу, которая будет массово сравнивать запросы с заданными шаблонами и искать несоответствия.

### **Triage & Troubleshooting**

В ходе тестирования от одной итерации к другой некоторые найденные проблемы могут повторяться. Особенно если дефект связан не с функциональной проблемой, а проблемой окружения или конфигурации. В таком случае решения проблем будут идентичными. Собирать данные о найденных проблемах можно в систему, которая будет являться некой базой знаний. А в случае нахождения уже существующей в системе проблемы можно будет легко найти способ ее решения.

В результате проделанного исследования были сформулированы предложения по совершенствованию процесса тестирования при условии совмещения автоматизации и ручного подхода. Данные рекомендации помогут равномерно распределять затраты на ручное тестирование и написание вспомогательных автоматизированных инструментов для тестирования. Основным преимуществом данного предложения является то, что автотесты или инструменты не заменяют процесс ручного тестирования, а лишь дополняют его, делают удобнее и проще. За счет этого можно добиться значительного уменьшения времени прохождения тестового сценария, а

также оптимизировать работу инженера-тестировщика и обеспечить тем самым высокое качество программного продукта.

Дальнейшее направление исследования заключается в моделировании процессов тестирования и сравнении выходных данных по выбранным критериям при использовании ручного тестирования и с использованием автоматизации, а также анализ эффективности использования того или иного метода тестирования.

### **Литература**

1. Золотухина Е.Б., Макарова Е.А., Беляков А.А. Обзор методов тестирования программного обеспечения // Аллея науки. 2018. Т. 4. № 6. С. 10-18.
2. Моисеев Д.А. Методология и процесс ручного тестирования // Надежность и качество сложных систем. 2017. № 3 (19).
3. Лопатина К.И., Темникова Е.А. Ручное и автоматизированное тестирование программного обеспечения отличия и взаимозаменяемость // Вестник современных исследований. 2017. № 10-1(13). С. 75–77.



**Маслюхин Сергей Михайлович**  
Университет ИТМО,  
факультет информационных технологий  
и программирования,  
аспирант группы № 7717,  
направление подготовки: 09.06.01 – Информатика  
и вычислительная техника,  
e-mail: sermasl2009@gmail.com

УДК 004.89

## **ОПТИМИЗАЦИЯ ГЛУБОКИХ НЕЙРОННЫХ СЕТЕЙ НА БАЗЕ ТРАНСФОРМЕРОВ**

**С.М. Маслюхин**

Работа выполнена в рамках темы НИР №620183 «Разработка виртуального диалогового помощника для поддержки проведения дистанционного экзамена на основе моделей-трансформаторов и понимания естественного и математического языка».

### **Аннотация**

В работе рассмотрено применение различных техник оптимизации нейронных сетей к глубоким нейронным сетям на базе трансформеров. Основное внимание уделено экспериментам в условиях ограниченных вычислительных ресурсов, доступных исследователям непромышленных организаций. Проведены эксперименты, отражающие эффективность применения базовых техник оптимизации.

### **Ключевые слова**

Нейронные сети, трансформеры, оптимизация, квантизация, прунинг, дистилляция, PyTorch, ONNX.

Развитие вычислительной техники, накопление больших объёмов данных и совершенствование алгоритмов машинного обучения позволили обучать глубокие и сверхглубокие нейронные сети. Наибольшее распространение в области обработки естественного языка в настоящее время получили глубокие нейронные сети на базе трансформеров [1]. С момента своего появления они удерживают лидирующие места на большей части текстовых задач, решаемых алгоритмами машинного обучения. Особенно хорошо модели трансформеров показали себя на задачах преобразования последовательности в последовательность, то есть таких задачах, в которых последовательности на входе в модель сопоставляется последовательность на выходе из модели. Широкому распространению моделей трансформеров способствовало развитие подхода переноса обучения с одной задачи на другую (transfer learning). Модели с большим количеством параметров, такие как GPT [2] и BERT [3], предобученные на больших объёмах текстовых данных без учителя, дообученные под конкретную задачу с учителем, показали лучшие на момент появления результаты на целом ряде текстовых задач. После появления первых предобученных моделей это направление начало активно развиваться и стало появляться много новых видов предобученных моделей. Со временем наблюдается стремительный рост размеров предобученных моделей, график тенденции роста размеров предобученных NLP моделей представлен на рис. 1.

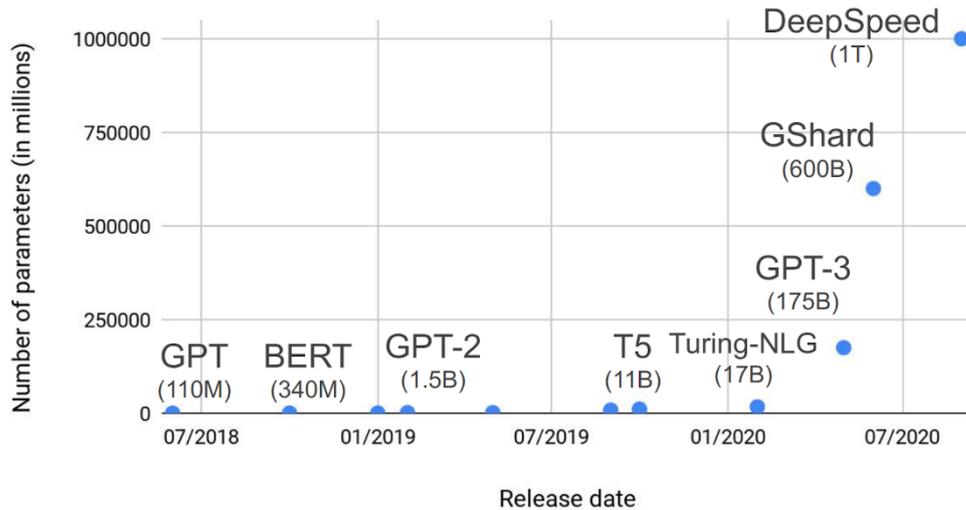


Рис. 1. Изменение размеров NLP моделей со временем

Практика применения моделей всё большего размера связана с тем, что модели с большим количеством параметров, обученные на большем количестве данных, достигают лучших результатов. Влияние размеров модели на значение функции потерь, а также на требуемое количество вычислительных ресурсов показано на рис. 2.

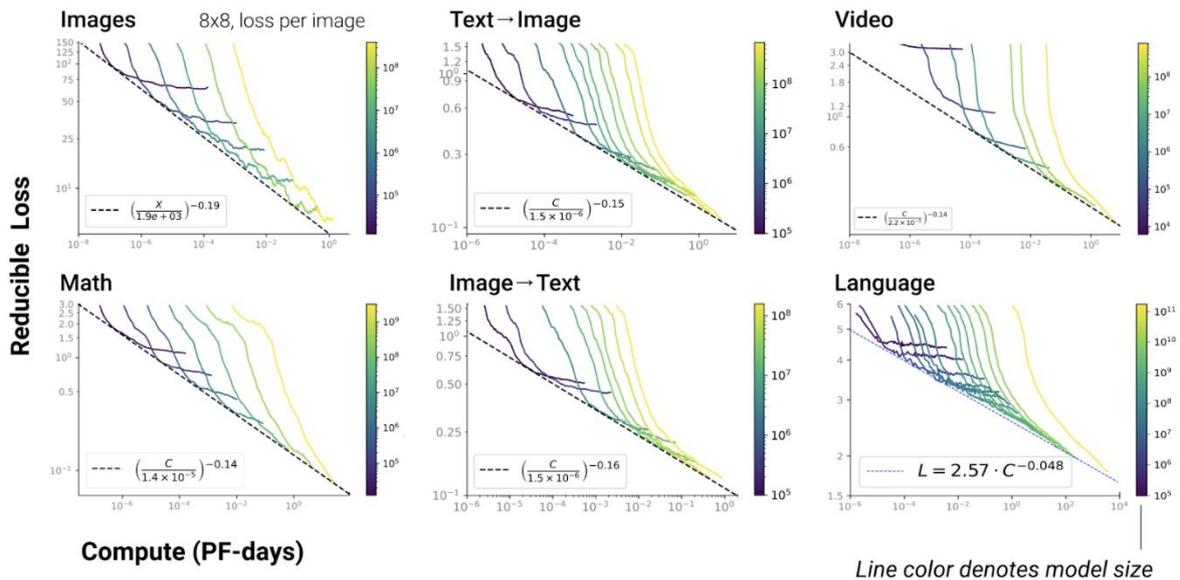


Рис. 2. Зависимость максимально достижимого качества модели и затрачиваемых вычислительных ресурсов от размера модели [4]

Таким образом, для достижения сравнимых результатов исследователи вынуждены работать с моделями всё большего размера. Обучение таких больших моделей становится очень дорогим и доступным только крупным корпорациям. Так, на предобучение модели GPT-3 по приблизительным оценкам было затрачено 4.6 млн долларов [5]. В этих условиях оптимизация глубоких нейронных моделей на базе трансформеров становится критически важной.

Существующие способы оптимизации нейронных сетей можно разделить на три группы: базовые, архитектурные и программные. К базовым оптимизациям относятся техники применимые к любым нейронным сетям. Такими техниками являются квантизация, прунинг и дистилляция. К архитектурным оптимизациям относятся

техники, связанные с изменением архитектуры трансформера, либо отдельных её компонентов. Оптимизациями такого типа являются замена вычислительно дорогого механизма self-attention на более эффективные аналоги, а также применение рекуррентных трансформерных архитектур. К программным оптимизациям можно отнести применение высокопроизводительных фреймворков глубокого обучения, таких как ONNX. В этой статье исследовались подходы к оптимизации, применимые к предобученным моделям. В первую очередь были рассмотрены квантизация и прунинг, а также проведено сравнение фреймворков PyTorch и ONNX на GPU и CPU. Эксперименты проводились на машине с CPU Intel Core i5-8400 и GPU GeForce GTX 1080 Ti. В качестве предобученной модели использовалась Conversational RuBERT, дообученная на задачу классификации на данных ru\_sentiment. В ходе эксперимента оценивалась производительность модели в режиме инференса, а также размер модели, занимаемый в памяти.

Квантизация нейросетевых моделей заключается в выполнении вычислений и хранении тензоров с меньшей точностью, чем точность чисел с плавающей запятой [6]. Существующие способы квантизации позволяют как конвертировать уже обученную модель в квантизованную, так и применить квантизацию в процессе обучения. Второй способ приводит к меньшим потерям в качестве классификации, чем первый. В табл. 1 представлены результаты применения квантизации к обученному чекпоинту.

Таблица 1

#### Результаты квантизации

| Device | Framework | Quantization | Inference performance | Size   | Accuracy |
|--------|-----------|--------------|-----------------------|--------|----------|
| CPU    | PyTorch   | No           | 13.04 it/s            | 711.5M | 77.42%   |
| CPU    | PyTorch   | Yes          | 21.57 it/s            | 454.9M | 78.41%   |
| CPU    | ONNX      | Yes          | 21.22 it/s            | 178.3M | 80.08%   |

По результатам эксперимента видно, что квантизация позволяет значительно сократить место, занимаемое моделью на диске, а также существенно увеличивает производительность модели в инференсе. При этом размер модели после квантизации фреймворком ONNX оказался заметно меньше, так как квантизация фреймворка ONNX применима к большему количеству типов слоёв нейросетевых моделей.

Прунинг нейросетевых моделей заключается в удалении наименее важных весов из сети [7]. В табл. 2 представлены результаты применения прунинга.

Таблица 2

#### Результаты прунинга

| Device | Framework | Pruning | Inference performance | Size   | Accuracy |
|--------|-----------|---------|-----------------------|--------|----------|
| CPU    | PyTorch   | No      | 13.04 it/s            | 711.5M | 77.42%   |
| CPU    | PyTorch   | Yes     | 13.07 it/s            | 711.5M | 64.91%   |

Как видно из таблицы, применение прунинга не влияет на скорость инференса и размер модели на диске. Качество классификации при этом значительно снижается. Это связано с тем, что операции на разреженных тензорах в данный момент не до конца реализованы существующими фреймворками, такими как PyTorch, поэтому прунинг применяется, но используется неэффективно. Фреймворк ONNX не поддерживает прунинг.

ONNX (Open Neural Network Exchange) – это высокопроизводительный фреймворк глубокого обучения. Он позволяет эффективно использовать вычислительные ресурсы системы, используя доступ к оптимизациям на аппаратном уровне. В табл. 3 представлены результаты сравнения фреймворков PyTorch и ONNX.

Таблица 3

Результаты сравнения PyTorch и ONNX

| Device | Framework | Inference performance | Size   | Accuracy |
|--------|-----------|-----------------------|--------|----------|
| CPU    | PyTorch   | 13.04 it/s            | 711.5M | 77.42%   |
| GPU    | PyTorch   | 109.06 it/s           | 711.5M | 77.42%   |
| CPU    | ONNX      | 17.40 it/s            | 711.5M | 81.12%   |
| GPU    | ONNX      | 179.89 it/s           | 711.5M | 81.12%   |

По результатам, представленным в таблице, видно, что переход к фреймворку ONNX позволяет получить значительный прирост в производительности, сохраняя при этом размер, занимаемый моделью на диске.

В ходе работы были рассмотрены различные способы оптимизации нейронных сетей на базе трансформеров. Полученные результаты позволяют сделать вывод об эффективности применения квантизации и использования фреймворка ONNX. Для применения прунинга необходима поддержка операций над разреженными тензорами на стороне фреймворка глубокого обучения.

В ходе эксперимента было рассмотрено влияние различных техник оптимизации на размер модели на диске и производительность в инференсе. Далее планируется исследование оптимизаций, применимых к процедуре обучения моделей.

## Литература

1. Vaswani A., Shazeer N., Parmar N., Uszkoreit J., Jones L., Gomez A.N., Kaiser L., Polosukhin I. «Attention Is All You Need», Advances in Neural Information Processing Systems 30. 2017. p. 5998-6008.
2. Improving Language Understanding with Unsupervised Learning [Электронный ресурс]. Режим доступа: <https://openai.com/blog/language-unsupervised/> (дата обращения: 23.02.2021).
3. Devlin J., Ming-Wei Chang, Lee K., Toutanova K., «BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding», In Proceedings of the 2019 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies, Volume 1, 2019. p. 4171–4186,
4. Henighan T., Kaplan J., Katz M., Chen M., Hesse C., Jackson J., Jun H., Brown T.B., Dhariwal P., Gray S., Hallacy C., Mann B., Radford A., Ramesh A., Ryder N., Ziegler D.M., Schulman J., Amodei D., McCandlish S. «Scaling Laws for Autoregressive Generative Modeling», ArXiv, abs/2010.14701, 2020.
5. OpenAI's GPT-3 Language Model: A Technical Overview [Электронный ресурс]. Режим доступа: <https://lambdalabs.com/blog/demystifying-gpt-3/> (дата обращения: 23.02.2021).
6. Dynamic quantization on BERT [Электронный ресурс]. Режим доступа: [https://pytorch.org/tutorials/intermediate/dynamic\\_quantization\\_bert\\_tutorial](https://pytorch.org/tutorials/intermediate/dynamic_quantization_bert_tutorial) (дата обращения: 23.02.2021).
7. Pruning tutorial [Электронный ресурс]. Режим доступа: [https://pytorch.org/tutorials/intermediate/pruning\\_tutorial](https://pytorch.org/tutorials/intermediate/pruning_tutorial) (дата обращения: 23.02.2021).



**Машина Екатерина Алексеевна**

Год рождения: 2000  
Университет ИТМО,  
факультет программной инженерии и компьютерной техники,  
студент группы №Р33113,  
направление подготовки: 09.03.04 – Системное и прикладное  
программное обеспечение,  
e-mail: mashina.katherina@gmail.com



**Балакшин Павел Валерьевич**

Год рождения: 1984  
Университет ИТМО,  
факультет программной инженерии и компьютерной техники,  
к.т.н.,  
e-mail: pvbalakshin@gmail.com

УДК 004.91

**УЧЕТ ОСОБЕННОСТЕЙ ПРЕОБРАЗОВАНИЯ  
УЗКОСПЕЦИАЛИЗИРОВАННЫХ ПРОФЕССИОНАЛЬНЫХ  
ТЕКСТОВ К ПОНЯТИЙНОМУ РЯДУ, ПОНЯТНОМУ  
НЕКВАЛИФИЦИРОВАННЫМ ПОЛЬЗОВАТЕЛЯМ, ПРИ ВЫБОРЕ  
ТЕХНОЛОГИЙ МАШИННОГО ПЕРЕВОДА, ОСУЩЕСТВЛЯЕМОГО  
ВНУТРИ ОДНОГО ЯЗЫКА**

**Е.А. Машина**

**Научный руководитель – к.т.н. П.В. Балакшин**

Работа выполнена в рамках инициативных исследовательских работ по проекту создания пациенто-ориентированной системы информационной поддержки медицинских волонтеров «PersonalMedHelper».

**Аннотация**

В работе рассматриваются особенности смысловой адаптации таких текстов к понятийному ряду неквалифицированных пользователей. На основе проведенного анализа делается аргументированный выбор технологий, выбираемых для создания автоматизированной системы такой трансформации, а также определяются направления дальнейших исследований и оценивается их практическая применимость.

**Ключевые слова**

Машинный перевод, основанный на примерах, рекуррентные нейросети, смысловая точность перевода, пользовательский тезаурус.

**Введение**

Современный этап научно-технической революции приводит к очень быстрому расширению понятийного ряда практически всех отраслей знаний, науки и технологий. Темп преобразований настолько высок, что эти изменения не успевают естественным образом интегрироваться в каждодневные языки общения. При этом на потребительский рынок выпускаются все более сложные образцы «пользовательских решений».

Все это приводит к тому, что детальное описание функционирования новых устройств все более усложняется для понимания конечного пользователя.

Это практически автоматически влечет за собой неправильное толкование покупателем новинок разнообразных «пользовательских инструкций». Соответственно, это приводит к повсеместному нежеланию пользователей детально знакомиться с подготовленными для них материалами. И в результате все это приводит к неполному использованию возможностей предложенных средств или даже к грубым ошибкам эксплуатации.

С другой стороны, открытость информационных ресурсов и высокопроизводительные средства телекоммуникации открывают конечным пользователям практически неограниченный доступ к узкоспециализированным текстам, которые дополнительно могут расширить знания и образовательный уровень непрофессиональных читателей. Однако и в этом случае отсутствие необходимой квалификации в ситуативно-интересующей предметной области не позволяет пользователю в полном объеме воспринять необходимую информацию.

### **Постановка задачи**

Очевидно, что решение задачи повышения понятности узкоспециализированного неквалифицированным пользователем можно проводить двумя путями:

— генерация специального контента, ориентированного на среднестатистического неквалифицированного пользователя, обладающего неким унифицированным объемом знаний (обычно таким уровнем считается сет остаточных знаний средней школы [1]),

— создание специализированных автоматизированных средств, позволяющих обеспечить неквалифицированному пользователю доступ к хорошо структурированной узкоспециализированным данным [2].

Естественно, наилучшим качеством адаптации специализированной информации для неквалифицированного пользователя обладает первый путь, сводящийся фактически к дублированию узкопрофессиональных текстов на повседневный язык неквалифицированных пользователей. Однако такой путь трансформации исходных профессиональных данных в понятный ряд неквалифицированных пользователей вряд ли можно считать экономически целесообразным (особенно принимая во внимание большое количество специализированной информации, которая генерируется различными отраслями знания сегодня).

В связи с этим актуальной видится задача создания автоматизированных механизмов адаптации профессиональных информационных ресурсов для нужд непрофессиональных пользователей.

По сути дела, эту задачу можно свести к созданию своеобразных переводчиков «с профессионального на обывательский». Причем следует иметь в виду, что наличие и обратного преобразования в подобном инструментарии тоже кажется необходимым (в том числе и для повышения адекватности обработки пользовательских запросов).

### **Особенности поставленной задачи, влияющие на метод ее решения**

В настоящее время задача поддержки непрофессионального пользователя в освоении им узкоспециализированных текстов в большинстве случаев решается путем использования многочисленных интерактивных словарей специальных терминов или разнообразных справочников [3]. Но это сегодня не может быть признано удовлетворительным, учитывая современную степень развития полностью автоматизированных средств лингвистического перевода [4].

Таким образом, ставится задача построения именно средства автоматической трансформации специализированных текстов для непрофессионалов, проводимой внутри одного языка.

На первый взгляд кажется, что решение задачи создания так называемой «вавилонской Золотой рыбки» [5] для перевода профессиональных текстов на язык непрофессионалов, проходящее фактически внутри одного языка, не так сложно, как создание переводчика между различными языками, имеющими, порой, совершенно иную структуру построения сообщений. И при ее решении можно воспользоваться лишь стандартными алгоритмами «интерлингвистических» переводчиков. Однако существует ряд причин, не позволяющих эффективно использовать технологии межязыкового автоматического перевода для адаптации специализированных текстов для нужд непрофессионалов, проводимой внутри одного языка.

К основным таким особенностям можно отнести следующее.

1. Во-первых, перевод специализированного текста для непрофессионала требует однозначности и точности интерпретации получаемого результата. В противном случае неквалифицированному пользователю нет необходимости пользоваться специализированными источниками для повышения своей информированности.

2. Во-вторых, узкопрофессиональные языковые системы оперируют понятийным рядом, в значительной части не имеющим однозначных аналогов в разговорных системах. Поэтому перевод специализированной терминологии требует для повышения степени однозначности множества сопутствующих уточнений, которые усложняют адаптированный материал. И это однозначно приводит к снижению понятности перевода.

3. В-третьих, узкопрофессиональные тексты (как правило, касающиеся наиболее передовых исследований и разработок) содержат большое количество специальных терминов, акронимов и аббревиатур, еще не получивших всеобщего признания. В связи с этим при переводе часто приходится составлять и использовать персонифицированные авторские тезаурусы.

4. В-четвертых, понятие «тезауруса непрофессионального читателя», описывающего понятийный ряд неквалифицированного пользователя (к которому и надо «приводить» путем упрощений и пояснений адаптируемый текст) на сегодняшний день представляется достаточно размытым и также требующим однозначного уточнения.

Очевидно, что в ближайшее время в связи с существенными различиями в структуре понятийных систем различных отраслей знания ожидать создания универсального решения для адаптации узкоспециализированных текстов практически невозможно [6]. Однако при создании подобных систем перевода все же представляется целесообразным использовать наиболее универсальные (но вместе с тем узкопрофессиональные) решения, позволяющие создать ограниченное число таких «профессиональных трансляторов» внутри одного языка.

### **Методы, которые предлагается использовать для решения задачи**

Анализ развития современных систем машинного перевода показывает, что возможно достаточно быстро создать средство адаптации профессиональных текстов, основываясь на широко используемых сегодня методах глубокого обучения (в том числе на основе рекуррентных нейросетей - Recurrent Neural Networks (RNN)), которые позволяют выстраивать варианты перевода на основе уже полученных данных) [7].

Однако прямое использование описанных выше методов перевода (лежащих в основе большинства современных автоматизированных транс-лингвистических переводчиков) практически неприменимо в случае переводов с «профессионального»

на «каждодневный невалифицированный» в связи с тем, что на сегодняшний день повсеместно отсутствуют систематизированные «зеркальные базы текстов» достаточного объема, на которых возможно проведение обучения нейросетей. А ведь известно, при небольших объемах сопоставимой информации говорить о правильности и устойчивости перевода, построенного на методах глубокого обучения, практически невозможно [8]. И это практически закрывает возможности использования методов обучения для построения систем адаптации узкоспециализированных текстов для невалифицированных пользователей, так как основными требованиями к подобному переводу являются его точность и однозначность [9]).

В связи с этим ставится задача определения иного подхода к построению автоматизированной системы адаптации специализированного контента.

Выбирая общую технологию автоматической адаптации профессионального медицинского контента для нужд непрофессионального пользователя, следует обратить внимание на то, что такие узкоспециализированные тексты хорошо структурированы. А использующиеся в них термины и описания действий большей частью стандартизованы в специализированные цепочки, часто называемые протоколами [10].

В связи с этим представляется целесообразным в первом приближении за основу технологии построения машинного перевода (с медицинского на каждодневный) взять Машинный Перевод Основанный на Примерах – Example-based Machine Translation (ЕВМТ) [11].

Рассматриваемый вид перевода может рассматриваться как перевод по аналогии, сводящийся к применению метода рассуждений на основе фрагментарных прецедентов к машинному обучению. Метод фактически сводится к тому, что в начале работы предложения в переводимом и исходном текстах разбиваются на определённые фрагменты-фразы, затем с помощью машинного обучения определяется соответствие фрагментов, после чего из переведенных таким образом фраз составляют предложение.

Результативность технологии ЕВМТ для адаптации узкоспециализированных текстов обуславливается тем, что она, с одной стороны, не требует обширных баз данных «зеркальных текстов», необходимых для обучения, с другой стороны, не требует делать и глубокий лингвистический анализ, необходимый для создания систем перевода, основанных на правилах.

При ЕВМТ предложения переводимого (адаптированного) текста разбиваются на определённые фразы, затем происходит перевод этих простейших фраз с учетом уже накопленного ранее опыта перевода. А после этого из этих фраз-фрагментов строится окончательный перевод предложения [11].

При этом технология ЕВМТ позволяет не проводить предварительное выстраивание всеобъемлющей системы правил и исключений, необходимых для построения фраз. Вместо этого в ЕВМТ при осуществлении перевода происходит использование фрагментов текстов, имеющих аналоги на языке, на который требуется перевод. Это позволяет говорить о ее универсальности для различных отраслей знания.

Причем следует отметить, что отдельные фрагменты алгоритмов перевода, основанного на примерах (правда, в неавтоматическом режиме), уже давно и успешно применяются в узкоспециализированных системах перевода отраслей знания, накладывающих на перевод особые требования к точности и однозначности [12].

Еще одним механизмом, призванным повысить производительность создаваемого решения для адаптации узкоспециализированного контента, основанного на методах и алгоритмах ЕВМТ, является технология Translation Memory (ТМ). Эта технология повышает скорость перевода за счет уменьшения объема проводимой работы при переводе повторяющихся фрагментов текстов [13].

При этом повышение производительности перевода обуславливается тем, что механизм ТМ базируется на сравнении фрагмента документа, подлежащего переводу, с данными, хранящимися в предварительно созданной базе переводов.

При осуществлении перевода система в общем массиве текста находит сегменты, которые уже были однажды переведены, и выбирает для текущего использования максимально похожие переводы из базы уже осуществленных переводов Translation Memory.

Таким образом, на основе знаний о предыдущих успешных переводах системой создается собственный специализированный словарь перевода фрагментов текста и терминов, постепенно дополняющий предварительно подгруженные словари терминов и соответствий, составленные независимо от опыта использования системы.

Однако следует иметь в виду, что сравнительная простота реализации системы перевода, основанного на технологии ЕВМТ, приводит к необходимости решения еще одной задачи в областях, где определяющими требованиями к системе перевода являются точность и однозначность.

Указанная проблема «смысловой точности перевода» обуславливается тем, что системы перевода, основанные на технологии ЕВМТ, не предусматривают проведение семантического анализа результирующего текста. В связи с этим без проведения дополнительных верификационных процедур доверять результату перевода можно лишь с заданной степенью вероятности.

Поэтому, чтобы обеспечить необходимое качество перевода при использовании технологии ЕВМТ, в систему необходимо дополнительно вводить специализированные механизмы верификации получаемого результата, предназначенные для оценки допустимости качества полученного результата перевода.

При этом одним из вариантов такой автоматической проверки может стать метод регулируемого изменения объемов словаря фраз и понятий, доступного для создаваемого автоматического переводчика. Сравнительная оценка устойчивости получаемого результата к объему предоставляемого словаря будет говорить о качестве перевода.

### **Выводы**

Рассмотренные выше особенности процесса трансформации узкоспециализированных медицинских текстов к виду, понятному непрофессиональным пользователям (который проводится внутри одного языка), позволяют говорить о том, что:

— наиболее эффективным механизмом для этого является Машинный Перевод Основанный на Примерах – Example-based Machine Translation (ЕВМТ), дополненный технологией Translation Memory (ТМ),

— при применении указанных технологий в качестве «исходных данных», на которых производится предварительное обучение системы, можно использовать не массивы «зеркальных текстов» (как правило, отсутствующие в необходимом объеме при переводе узкопрофессиональных текстов), а толковые словари специализированных терминов,

— для обеспечения необходимого качества перевода в создаваемую систему необходимо дополнительно вводить специализированные механизмы верификации, которые могут быть основаны на методе регулируемого изменения объемов исходных словарей,

— наиболее перспективным направлением внедрения подобных систем автоматической генерации адаптированного для неквалифицированного пользователя узкоспециализированного контента является телемедицина в той ее части, которая относится к созданию пациенто-ориентированных информационных систем.

## Литература

1. Машина Е.А., Бейлин М.Т. Анализ нормативов описывающих оказание первичной доврачебной помощи для определения потребностей муниципальных медицинских волонтеров, работающих с применением систем информационной поддержки // Сборник трудов VIII конгресса молодых ученых (Санкт-Петербург, 15-19апреля 2019г.). 2019. Т. 3. С. 264-266.
2. Машина Е.А., Нелепко Л.Н. Создание пациенто-центричной телемедицинской системы PersonalMedHelper в открытом междуниверситетском пространстве коворкинга // Труд молодежи в XXI веке: вызовы и возможности. Диалог поколений. Материалы Петербургского международного молодежного форума труда - 2019 (Санкт-Петербург, 27–28февраля 2019г.). 2019. С. 243-257.
3. Нелюбин Л.Л. Введение в технику перевода (когнитивный теоретико-прагматичный аспект) // Нелюбин Л.Л. 5-е изд., стер. М.: ФЛИНТА. 2016. 216 с.
4. Демьянков В.З. Когнитивная лингвистика как разновидность интерпретирующего подхода. / Демьянков В.З. // Вопросы языкознания. 1994. № 4 (35). С. 17-33.
5. Somers H. Computers and Translation: A Translator's Guide. John Benjaminus Publ. Company. Amsterdam. 2003. 351 p.
6. Bowker L. Computer-Aided Translation Technology: A Practical Introduction – University of Ottawa Press. Ottawa. 2002. 184 p.
7. Машина Е.А., Баркалов М.М. Возможности использования методов машинного обучения для анализа массивов фактографических данных, генерируемых при использовании пациенто-ориентированных телемедицинских систем // Сборник трудов VIII конгресса молодых ученых (Санкт-Петербург, 15-19апреля 2019г.). 2019. Т. 3. С. 260-263.
8. Rosch E.H. Cognitive Representations of Semantic Categories // Journal of Experimental Psychology: General, 1975. Vol. 104. № 3, p. 192– 233.
9. Машина Е.А. Особенности создания пациенто-ориентированных телемедицинских систем // Сборник трудов VIII конгресса молодых ученых (Санкт-Петербург, 15-19 апреля 2019г.). 2019. Т. 3. С.256-259.
10. Машина Е.А., Нелепко Н.Н. Структура и источники специализированных данных системы информационной поддержки муниципальных волонтеров, оказывающих доврачебную медицинскую помощь // Сборник трудов VIII конгресса молодых ученых (Санкт-Петербург, 15-19 апреля 2019г.). 2019. Т. 3. С. 267-270.
11. Somers H. Example-based Machine Translation // Machine Translation, 1999. № 14, p. 113-157.
12. Борисов В.В. Аббревиация и акронимия. Военные и научно-технические сокращения в иностранных языках / Борисов В.В.; под ред. Швейцера А.Д. М: Военное издательство. 2004. 320 с.
13. Kit C., Pan H., Webster J.J. Example-based Machine Translation: A New Paradigm // Translation and Information Technology / Chan S.W. (ed.). Hong Kong: Chinese University of Hong Kong, 2002. p. 57-78.



**Мишина Маргарита Эдуардовна**

Год рождения: 1998  
Университет ИТМО,  
Институт дизайна и урбанистики,  
студент группы № С41041,  
управление подготовки: 09.04.03 – Прикладная информатика.  
email: mishina.rita22@yandex.ru



**Гурина Екатерина Павловна**

Год рождения: 1981  
Университет ИТМО,  
Институт дизайна и урбанистики,  
преподаватель практики.  
ООО «Интеллектуальные медиа системы»,  
руководитель группы,  
email: epgurina@gmail.com

УДК 004.942

**АНАЛИЗ ИНСТРУМЕНТОВ ПОИСКА  
ОПТИМАЛЬНОГО МАРШРУТА  
В ТЕЛЕКОММУНИКАЦИОННОЙ СЕТИ В СРЕДЕ ARCGIS**

**М.Э. Мишина, Е.П. Гурина**

**Научный руководитель – преподаватель практики Е.П. Гурина**

Работа выполнена совместно с ООО «Интеллектуальные медиа системы» в рамках магистерского исследования «Разработка алгоритма программного модуля по расчету вариантов прокладки телекоммуникационной сети с учетом существующей инфраструктуры в рамках ГИС оператора связи» по программе обучения «Цифровые технологии умного города».

**Аннотация**

В работе рассмотрены основные возможности и ограничения моделей представления сетевого набора данных Network Dataset и Geometric Network, реализованных в геоинформационной системе ArcGIS. Полученные результаты сравнительного анализа двух моделей доказывают допустимость использования инструментов Network Dataset и Geometric Network при решении задачи оптимизации маршрута проектируемой линии связи в телекоммуникационной сети.

**Ключевые слова**

Телекоммуникации, оператор связи, сквозная автоматизация бизнес-процессов, проектирование линии связи, мультимодальная сеть, ГИС-платформа, геоинформационные системы, модели представления данных.

В наши дни цифровизация меняет все привычные индустрии и модели их развития. Традиционный бизнес телекоммуникационных компаний постепенно уходит в прошлое, ставя перед компаниями новые задачи и открывая возможности для роста. Будущее телеком-отрасли, связанное с глобальными вызовами современности

и новыми изменениями в законодательстве, активно обсуждается экспертами, представителями власти и бизнеса на научно-практических конференциях, форумах и пленарных заседаниях.

Сегодня для удержания позиций на рынке операторам связи приходится совершенствовать технологии, подстраиваться под изменяющиеся потребности клиентов, предлагать новые товары и услуги, а также расширять сферу деятельности. В связи с этим все чаще встает вопрос о необходимости вложения инвестиций в развитие информационных систем, обеспечивающих сквозную автоматизацию бизнес-процессов таким образом, чтобы находить как можно больше новых клиентов и повышать лояльность среди уже имеющихся. Вместе с тем с увеличением числа пользователей услуг и развитием телекоммуникационной сети у оператора связи закономерно возникает потребность в разработке современных решений по расчету оптимальных вариантов прокладки новых линий связи до конечных потребителей.

Одним из возможных решений задачи оптимизации маршрута проектируемой линии связи является трассировка мультимодальной сети с учетом весовых критериев. В данном случае мультимодальность сети заключается в объединении существующей кабельной инфраструктуры оператора и коммуникаций, вдоль которых возможна прокладка новых участков линии связи (туннелей телефонных канализаций, линий ЛЭП, пролетов между домами и т. д.). Иными словами, мультимодальная сеть представляет совокупность всех потенциально возможных вариантов маршрутов линий связи (как существующих, так и проектируемых) от всякой точки А до всякой точки Б на сети. Трассировка такой сети, осуществляемая с учетом весовых критериев, предполагает минимизацию основного критерия (импеданса), а также расчет дополнительных весов, выражающих ресурсные ограничения кабельных маршрутов (количество необходимых свободных волокон, транзитных соединений и т. д.) [1].

Моделирование сетей, представленных большим количеством объектов, имеющих пространственную привязку и семантическую информацию, их характеризующую, целесообразно производить на базе геоинформационных систем (ГИС), отличающихся высокой гибкостью в обработке и управлении гетерогенными данными [2]. Так, в ГИС ArcGIS моделирование мультимодальной сети осуществляется при формировании сетевого набора данных, представленного геометрическими примитивами и основанного на математическом аппарате теории графов. На данный момент в концепции ArcGIS реализовано три модели представления сетевого набора данных, отличающихся аналитическими возможностями и ограничениями – Network Dataset, Geometric Network, Utility Network. Однако, поскольку модель Utility Network поддерживается исключительно последними версиями программного обеспечения, в дальнейшем будут рассмотрены наиболее доступные подходы к моделированию сети – Network Dataset и Geometric Network.

### **Модель Network Dataset**

Модель представления сетевых данных Network Dataset, как правило, используется для представления сети в виде неориентированного графа. Данная модель получила широкое распространение в решении задач, связанных с анализом транспортных систем. Однако авторы некоторых исследований также обосновывают применение Network Dataset для моделирования инженерных сетей [3]. Так, например, неориентированный граф хорошо иллюстрирует возможность организации двухсторонних соединений в сетях телекоммуникаций.

Модель Network Dataset состоит из сетевых элементов – ребер, соединений, поворотов. Источниками сетевых элементов являются линейные и точечные классы пространственных объектов, сформированные в базе геоданных оператора связи. Связь сетевых элементов с объектами-источниками хранится в виде автоматически

сформированной логической сети – набора системных таблиц базы геоданных, аналогичных матрицам смежности и связности, используемым для решения задач на графах (рис. 1). При этом внесение каких-либо изменений в исходные классы пространственных объектов требует «ручного» обновления набора сетевых элементов для актуализации построенной сети.

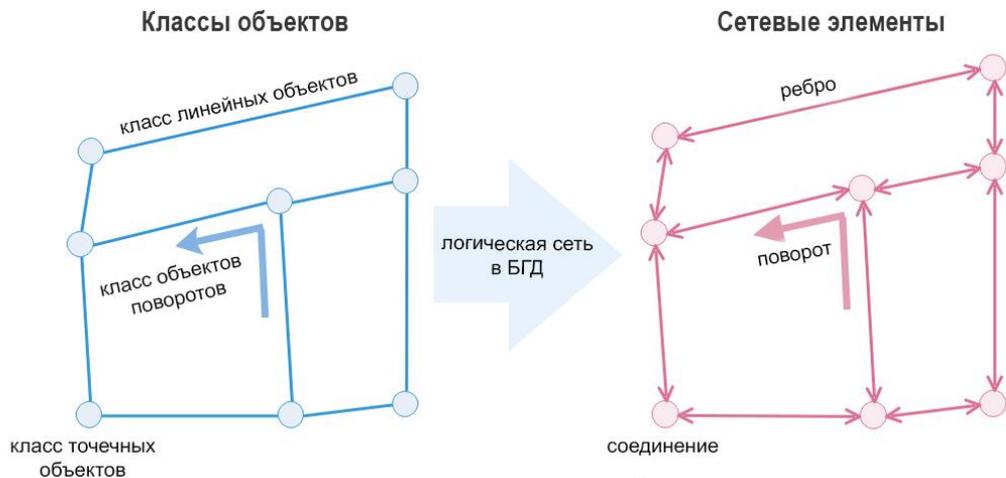


Рис. 1. Классы объектов и сетевые элементы модели Network Dataset

При формировании логической сети в рассматриваемой модели наравне с созданием сетевых элементов осуществляется настройка правил связности, основанных на наличии топологически корректных вершин сопряжения линейных и точечных объектов-источников. Возможность определения различных групп связности относительно исходных объектов в Network Dataset позволяет автоматически моделировать мультимодальную сеть, сохраняя целостность и обособленность нескольких типовых подсетей. При этом связность в мультимодальной сети позволяет учитывать данные о высоте сетевых элементов и не создавать лишние узлы на мнимом пересечении ребер (рис. 2).

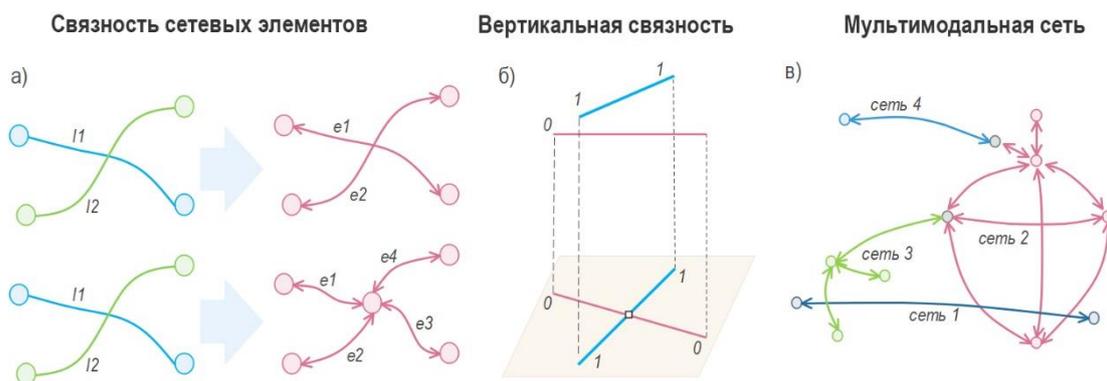


Рис. 2. Классы объектов и сетевые элементы модели Network Dataset

В данной модели для проектирования новых участков сетевой инфраструктуры используются механизмы расчета Route и Closest Facility, основанные на алгоритме Дейкстры. Метод расчета Route обеспечивает возможность построения кабельного маршрута между двумя точками, указанными на сети, тогда как метод Closest Facility рассчитывает путь от указанной точки до ближайшего транзитного соединения.

В каждом из методов поиск подходящего маршрута осуществляется с указанием сетевых атрибутов: импеданса как основного веса (например, стоимость строительства линии связи) и ограничений в виде логических условий к вариативным параметрам, выступающим в качестве признаков элементов сети (например, необходимое количество свободных оптических волокон). Оба метода расчета модели успешно реализуют поиск оптимального варианта прокладки нового участка линии связи за счет минимизации накопленного значения импеданса. Однако следует отметить присущую методам модели Network Dataset излишнюю функциональность, ориентированную на решение транспортных задач [4].

### Модель Geometric Network

Модель Geometric Network является наиболее распространенным способом представления сетевых данных инженерных систем и сетей с направленным движением ресурсов. Основная сфера применения данной модели определяет некоторые особенности ее использования при выполнении различных типов сетевого анализа.

Аналогично Network Dataset, моделирование сетевых элементов, ребер и соединений в Geometric Network осуществляется на основе классов линейных и точечных объектов, хранящихся в базе геоданных (рис. 3). Однако в данном случае логическая сеть, определяющая взаимосвязь между сетевыми элементами и объектами-источниками в виде набора системных таблиц, динамична. Это позволяет автоматически, без вмешательства пользователя, осуществлять пересчет геометрических и атрибутивных характеристик сети при внесении изменений в исходные классы объектов.



Рис. 3. Классы объектов и сетевые элементы модели Geometric Network

В модели Geometric Network отсутствует прямая возможность формирования мультимодальной сети. Однако при настройке правил связности сетевых элементов учитываются типы и подтипы объектов сети, которые могут иметь связь друг с другом, а также кардинальность такой связи. Кроме того, при сопряжении сетевых элементов в рассматриваемой модели используется встроенная топология базы геоданных, позволяющая на этапе построения сети автоматически идентифицировать и устранять некорректные пространственные отношения исходных объектов (рис. 4).



Рис. 4. Связность сетевых элементов в Geometric Network

Поскольку инженерным сетям (водопроводам, нефтепроводам и др.) зачастую свойственно строго ориентированное перемещение вещества от точки А к точке Б, Geometric Network позволяет моделировать сеть в виде ориентированного графа, определяя направление потока через оцифровку ребер или посредством назначения объектов-источников и объектов-приемников. Несмотря на то, что поиск оптимального кабельного маршрута не требует указания ориентации ребер в мультимодальной сети, данная функциональная возможность потенциально может быть полезной для оператора связи при осуществлении поиска абонентов, пострадавших из-за аварии на узле связи. Для этого предусмотрены механизмы расчета Trace Downstream и Trace Upstream.

Проектирование новых участков сетевой инфраструктуры в модели Geometric Network осуществляется с помощью метода трассировки Find Path, использующего веса и весовые фильтры для минимизации импеданса (например, стоимости прокладываемого кабельного маршрута) и соблюдения ресурсных ограничений сети (например, необходимое количество свободных оптических волокон). Метод Find Path осуществляет поиск наилучшего варианта пути между двумя точками, установленными на соединениях или ребрах сети. Однако проектирование маршрута от точки до ближайшего транзитного соединения потребует выполнения предварительных операций геообработки над объектами соединений сети [4].

### Заключение

Проведенный сравнительный анализ моделей представления сетевого набора данных в ArcGIS позволил определить основные возможности и ограничения при использовании моделей Network Dataset и Geometric Network в качестве ГИС-инструментов поиска оптимального маршрута. Исходя из полученных результатов, следует сделать вывод о допустимости использования моделей при решении задачи оптимизации маршрута проектируемой линии связи в телекоммуникационной сети. Каждая из реализованных моделей имеет ряд упомянутых преимуществ и недостатков, требующих отдельного рассмотрения в каждом конкретном случае. Тем не менее, при реализации автоматизированных решений в корпоративных информационных системах рекомендуется отдавать предпочтение модели Geometric Network, отличающейся наличием встроенной поддержки топологии и автоматическим обновлением системных таблиц базы геоданных при внесении изменений в исходные классы объектов сети. Использование данной модели позволяет снизить влияние человеческого фактора, а также сделать процесс управления данными удобным и доступным для пользователя.

### Литература

1. Солдатенко А.А. Алгоритм оптимальной маршрутизации в мультисервисных телекоммуникационных сетях. // Прикладная дискретная математика. Приложение. 2018. № 11. С. 122-127.
2. Щербинин М.В. Разработка и исследование принципов и методики построения информационно-телекоммуникационных систем на базе ГИС-технологий: автореф. дис. ... канд. техн. наук / Щербинин М.В. М. 2007. 24 с.
3. Карпачевский А.М., Новаковский Б.А. Возможности использования инструментов ГИС-пакета ArcGIS для структурного анализа электрических сетей. // Геоинформатика. 2019. № 2. С. 1-11
4. Zeiler Michael. Modeling our world: the Esri guide to geodatabase concepts. – 2 ed. 2010. 308 p.



**Назаренко Ульяна Кирилловна**

Год рождения: 1999  
Университет ИТМО,  
факультет инфокоммуникационных технологий,  
студент группы № К3443,  
направление подготовки: 45.03.04 – Интеллектуальные  
системы в гуманитарной сфере,  
e-mail: uliana\_nazarenko@bk.ru



**Козырева Алена Игоревна**

Год рождения: 1999  
Университет ИТМО,  
факультет инфокоммуникационных технологий,  
студент группы № К3443,  
направление подготовки: 45.03.04 – Интеллектуальные  
системы в гуманитарной сфере,  
e-mail: alyonakozyr@gmail.com

УДК 004.048

**ИНТЕГРАЦИЯ ПЛАТФОРМ СОЦИАЛЬНЫХ МЕДИА  
И СПЕЦИАЛИЗИРОВАННЫХ ВЕБ-РЕСУРСОВ  
ДЛЯ ЭФФЕКТИВНОГО ИСПОЛЬЗОВАНИЯ  
ВЫСОКОТЕХНОЛОГИЧНОЙ МЕДИЦИНСКОЙ ИНФОРМАЦИИ**

**У.К. Назаренко, А.И. Козырева**

**Научный руководитель –к.т.н. с.н.с. Н.Ф. Гусарова**

Работа выполнена в рамках темы НИР №620178 «Искусственный интеллект в медицине».

**Аннотация**

В работе рассмотрена проблема интеграции платформ социальных сетей и специализированных веб-ресурсов для эффективного использования высокотехнологичной медицинской информации. Изучена структура организации и самоорганизация участников сообществ в социальных сетях, направленных на эффективное использование высокотехнологичной медицинской информации, проведен проблемно-ориентированный анализ потребностей и мотиваций участников этих сообществ. На основе анализа разработано решение – специализированные чат-боты, обеспечивающее интеграцию платформ социальных сетей и специализированных веб-ресурсов для эффективного использования высокотехнологичной медицинской информации.

**Ключевые слова**

Чат-бот, здравоохранение, высокотехнологичная медицинская информация, структура интернет-сообщества, общедоступный ресурс, Телеграм-каналы, сообщество ВКонтакте.

В современном мире профилактика и лечение самых смертоносных заболеваний во всем мире все больше полагаются на высокие медицинские технологии [1]. Они, в

свою очередь, основаны на получении и использовании высокотехнологичной медицинской информации, где получение информации связано с наличием соответствующего оборудования, а ее использование для диагностики и лечения – с возможностью ее правильной интерпретации. Для интерпретации своей высокотехнологичной медицинской информации в условиях отсутствия врачей и недоверия к ним, что особенно характерно для развивающихся стран, люди пытаются использовать доступные Интернет-источники и коллективный сетевой интеллект, то есть апеллируют к коллективному мнению [2, 3]. Таким образом, возникает проблема интеграции платформ социальных сетей и специализированных веб-ресурсов для эффективного использования высокотехнологичной медицинской информации.

В работе экспериментально изучена структура общедоступных медицинских интернет-сообществ, характерных для России, и обнаружено, что им свойственна самоорганизация. Для формирования мотивационной составляющей разработанного веб-ресурса на базе Университета ИТМО по эффективному использованию медицинской информации используется выявленная структура уже существующих сетевых сообществ медицинской направленности. Предлагается использование специализированных чат-ботов как эффективное средство интеграции развитого ресурсного и сетевого сообществ.

Интернет-сервисы, с помощью которых опытные врачи предоставляют пациенту отчеты о его высокотехнологичной медицинской информации, представлены в странах с развитой медициной. Эти сервисы представляют собой веб-сайты или приложения со строго определенной бизнес-моделью, которая подразумевает законность и безопасность передаваемой информации. Врачи в этих службах – в основном ведущие специалисты, и цена их услуг соответствующая, что доступно не всем. Аналогичные сервисы существуют и в России, однако таких сервисов очень мало, они платные, и у них очень узкая специализация.

Проведенный анализ позволил сделать некоторые выводы. Использование сообществ в социальных сетях удобно и доступно всем слоям населения, но при этом не гарантирует легитимность и достоверность полученной интерпретации. В таких сообществах существует «спонтанный краудсорсинг», поддерживаемый самоорганизацией, но его использование в качестве механизма мотивации не изучено. С другой стороны, существующие специализированные медицинские услуги обладают высоким уровнем легитимности и надежности, но большинство из них дорогие и не предусматривают использование входящей информации в качестве источника для формирования наборов данных для использования в искусственном интеллекте.

В работе анализируются два медицинских сообщества проекта «ВКонтакте» по кардиологии, два канала по кардиологии и два канала по радиологии мессенджера «Telegram».

Анализ медицинских сообществ показал активный рост подписчиков с декабря 2019 года. Заметен рост общего количества просмотров для всех публикаций, что указывает на повышение уровня интереса людей к данным медицинским тематикам (в качестве примера – рис. 1-3).

Для каждого сообщества ВКонтакте были рассчитаны 22 переменные: характеризующие пользователей сообщества (переменные 1-9), «сети дружбы» (переменные 10-18) и структурные показатели пользователей сообщества (переменные 19-22). Описание переменных представлено в табл. 1.

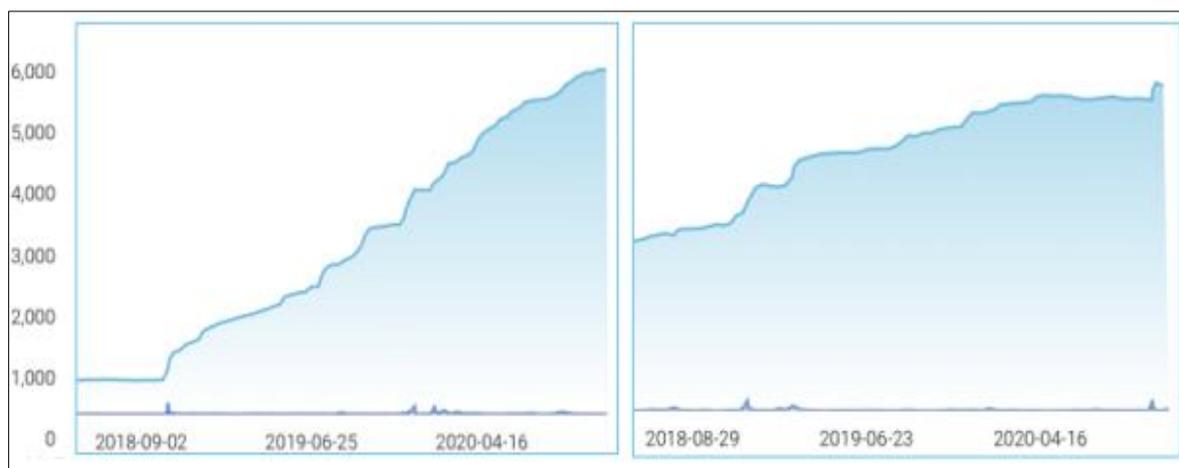


Рис. 1. Статистика по количеству подписчиков каналов по кардиологии мессенджера «Telegram»

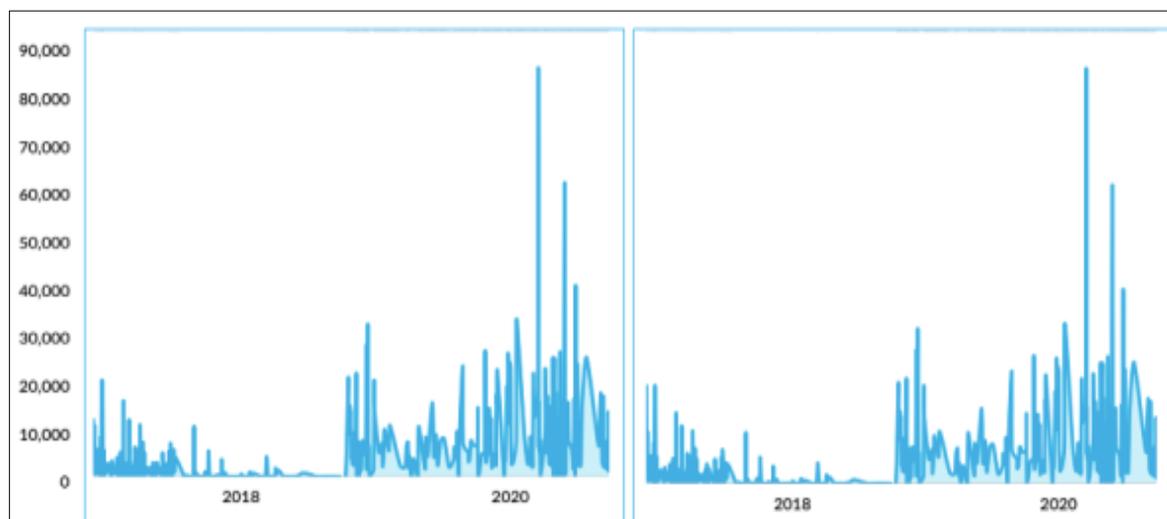


Рис. 2. Статистика по общему количеству просмотров для всех публикаций каналов по кардиологии мессенджера «Telegram»

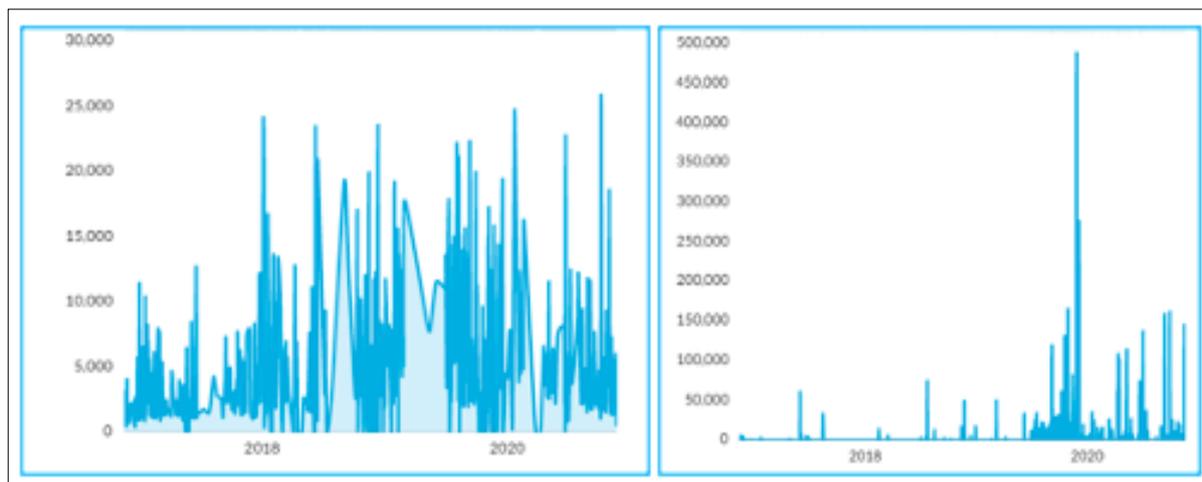


Рис. 3. Статистика по общему количеству просмотров для всех публикаций каналов по радиологии мессенджера «Telegram»

**Описание посчитанных переменных**

| №  | Переменная               | Описание   |
|----|--------------------------|--|
| 1  | Active Users             | Число «активных» участников группы среди чистой полной численности группы. Под активностью понимается совершение хотя бы одного из возможных онлайн-действий в группе: написание поста, написание комментария, написание сообщения в обсуждениях, оставление «лайка» |
| 2  | Population               | Полная численность онлайн-группы, включающая число формальных членов группы (вступивших в группу) и всех пользователей, проявивших некую коммуникационную активность на страницах группы   |
| 3  | Clear population         | Чистая полная численность онлайн-группы без удаленных/забаненных пользовательских аккаунтов  |
| 4  | Members                  | Число формальных членов группы среди чистой полной численности группы  |
| 5  | Share Active Members     | Доля активных участников, одновременно являющихся формальными членами группы   |
| 6  | Connected Users          | Численность связанных пользователей внутри онлайн-группы, имеющих хотя бы 1 «дружбу» с другим участником группы  |
| 7  | Connected Users Share    | Доля связанных пользователей внутри онлайн-группы, имеющих хотя бы 1 «дружбу» с другим участником группы   |
| 8  | Isolates                 | Число изолятов (вычисляется как разность между чистой полной численностью группы и связанными пользователями)  |
| 9  | Isolates Share           | Доля изолятов от чистой полной численности группы  |
| 10 | Edges                    | Число ребер сети «дружбы»  |
| 11 | Connected Components     | Доля вершин в максимальной связанной компоненте сети «дружбы»  |
| 12 | Vertex Giant Component   | Плотность сети «дружбы». Плотность сети – это отношение существующих связей графа ко всем возможным для графа с таким же числом вершин   |
| 13 | Density                  | Плотность сети «дружбы» без изолятов (т. е. только среди связанных пользователей)  |
| 14 | Modularity (from 0 to 1) | Значение модулярности (от 0 до 1), показывающее насколько выражена и кластерная структура в графе, насколько легко граф поддается разбиению на подгруппы вершин  |
| 15 | Clusters                 | Число выявленных кластеров для сети без изолятов   |
| 16 | Mean Geodesics           | Средняя геодезическая дистанция сети «дружбы». Геодезическая дистанция – кратчайший путь между любой парой узлов сети  |
| 17 | Diameter                 | Диаметр сети «дружбы» (максимальная геодезическая дистанция в графе)   |
| 18 | Mean Degree              | Среднее значение степени центральностей сети «дружбы» для полной сети. Степень центральности (degree centrality), $CD(i)$ – характеристика вершины, равная количеству связей одной вершины с другими вершинами [3]   |
| 19 | Female Share             | Доля пользователей женского пола среди чистой полной численности группы  |
| 20 | Writer Share             | Доля участников, создающих контент в группе в виде постов или комментариев   |
| 21 | Liker Share              | Доля участников, не создающих контент в группе, но ставящих «лайки» к контенту группы  |
| 22 | Passive Share            | Доля пассивных участников, не «лайкающих» и не создающих контент. Являются подмножеством формальных членов группы  |

По полученным структурным показателям «сети дружбы» сообщества видим, что данная сеть имеет большой диаметр, агенты сети слабо связаны, выделяются определенные кластеры и присутствует большое число изолятов (рис. 4).

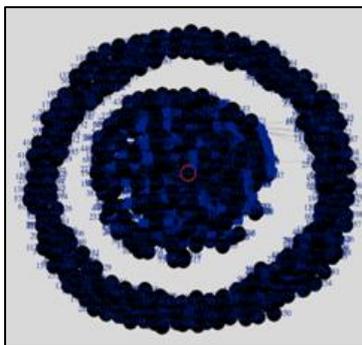


Рис. 4. Общая структура сети одного из сообществ

Построенные «сети дружбы» состоят из множества подграфов, что указывает на то, что агенты сети слабо связаны между собой, то есть можно предположить, что участники данных сообществ не преследуют цели знакомства друг с другом и заведению новых друзей, а скорее единично взаимодействуют с некими «лидерами».

Для получения эффективной оценки высокотехнологичной медицинской информации жители России обращаются преимущественно в публичные Интернет-пространства. Возникают и становятся популярными сообщества, работающие в режиме «спонтанного краудсорсинга», они имеют ярко выраженную самоорганизацию. Пользователи разделены на группы – пациенты, студенты и врачи – со своими интересами, которые не противоречат друг другу и там полностью удовлетворены. Это позволяет таким сообществам бесплатно предоставлять необходимые информационные услуги своим участникам. Медицинская информация передается и распространяется открыто, без учета требований конфиденциальности и законных требований. Это указывает на актуальность создания общедоступного решения для оценки высокотехнологичной информации, сочетающего доступность и простоту общения, присущих социальным сетям, предоставляя пациентам законное второе мнение и в то же время позволяя использовать предоставленную информацию для формирования наборов данных. В работе предлагается решение как комбинация специализированного сетевого ресурса и набора чат-ботов, которые интегрируют его с сетевыми платформами.

Полученные результаты позволили сформулировать требования к предлагаемому решению:

1. Сообщества ВКонтакте и Telegram-каналы недостаточно легитимны. Чат-бот способен решить эту проблему, поскольку интегрируется с высокотехнологичным сервисом.

2. Люди должны быть уверены в том, что сервису можно доверять. Поэтому при разработке чат-бота необходимо обеспечить сокрытие личных данных.

3. Для удобства пользователям необходима возможность предоставлять данные в различных графических форматах.

4. Необходима возможность обмена информацией в простой диалоговой форме. Важно разработать разные сценарии диалогов, разные уровни описания для разных групп пользователей. Поэтому необходимо предусмотреть несколько чат-ботов, ориентированных на определенные группы пользователей.

В настоящее время на платформе Университета ИТМО запущен специализированный интернет-ресурс по эффективному использованию

высокотехнологичной медицинской информации «Медицинская краудсорсинговая платформа Университета ИТМО». Архитектура ресурса ИТМО-МСП представляет собой реализацию модульного веб-сервиса с хранением данных в виде частной цепочки блоков (рис. 5). Ресурс работает по принципу телерадиологии. Его наборы данных формируются благодаря лечащим врачам и их пациентам, которые готовы предоставить свою анонимную медицинскую информацию. Платформа обладает высоким уровнем конфиденциальности и безопасности. Ресурс позволяет использовать высокотехнологичные инструменты: это послужило мотивацией для создания бота.

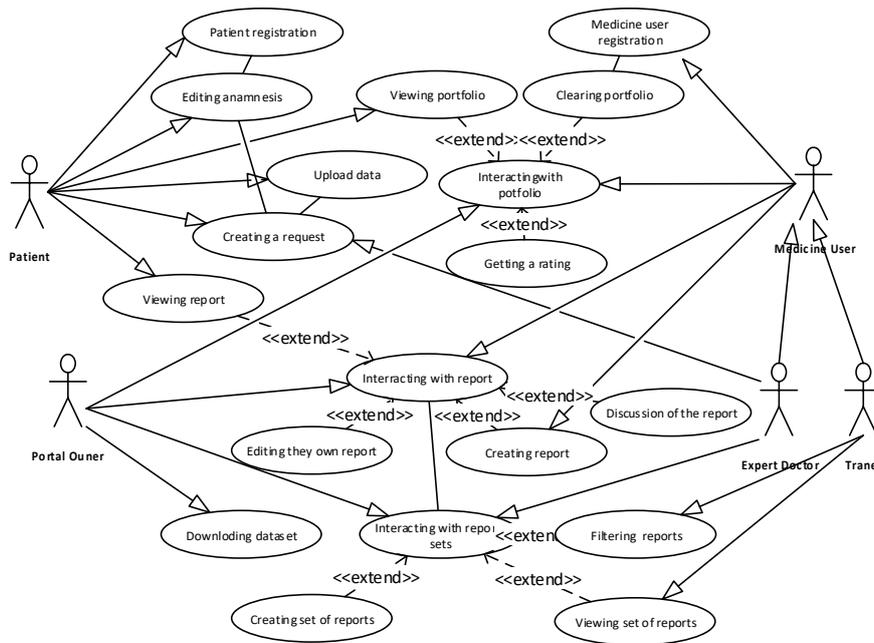


Рис. 5. Диаграмма использования Медицинской краудсорсинговой платформы Университета ИТМО (ИТМО-МСП) (по Пчелкину А.Ю.)

Чат-бот служит инструментом интеграции (рис. 6). Благодаря внедрению в социальные сети, он вовлекает людей в использование высокотехнологичного инструмента и объединяет их, тем самым удовлетворяя их потребности. Чат-бот прост в использовании и не требует специальных знаний.

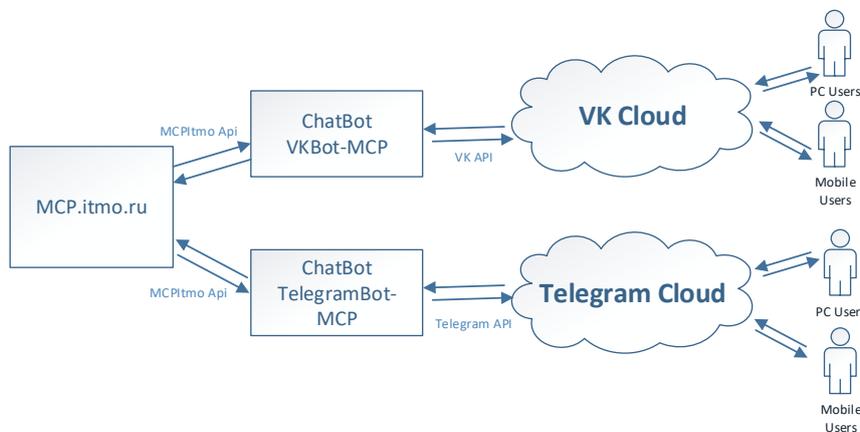


Рис. 6. Диаграмма использования чат-ботов

Взаимодействие пользователей социальных платформ с сервисом осуществляется через механизм чат-ботов. Предложенные чат-боты (TelegramBot-MCP

и VKBot-МСП) разрабатываются индивидуально для каждой платформы, поскольку механизм взаимодействия чат-бота с социальной платформой зависит от API, предлагаемого платформой.

Бизнес-логика взаимодействия пользователя с ботом реализована на стороне программного модуля бота. Социальные платформы предоставляют пользователям диалоговый интерфейс с чат-ботом в знакомой среде. Через чат-боты пользователь может отправить платформе запрос на получение второго мнения и получить ответ, как только он будет готов.

Взаимодействие чат-ботов и основной части платформы осуществляется через специальный API, позволяющий масштабировать систему и прозрачно подключать другие социальные платформы.

В работе была рассмотрена возможность создания общедоступного ресурса для эффективного использования высокотехнологичной медицинской информации, исследовали структуру организации участников сообщества в социальных сетях, а также проанализировали публичные медицинские интернет-пространства. Поставленные нами задачи успешно выполнены.

По результатам анализа медицинских сообществ предложено решение, обеспечивающее интеграцию платформ социальных сетей и специализированных веб-ресурсов для эффективного использования высокотехнологичной медицинской информации.

### Литература

1. Ahuja A.S. The impact of artificial intelligence in medicine on the future role of the physician. PeerJ. 2019; 7: e7702. Published online 2019 Oct 4. doi: 10.7717/peerj.7702 PMID: PMC6779111 [Электронный ресурс]. Режим доступа: <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC6779111/> (дата обращения: 22.11.20).
2. Allyn J. International Radiology Societies Tackle Radiologist Shortage. Feb. 20, 2020. [Электронный ресурс]. Режим доступа: <https://www.rsna.org/en/news/2020/February/International-Radiology-Societies-And-Shortage> (дата обращения: 22.11.20).
3. Freeman L. C. The Development of Social Network Analysis: A Study in the Sociology of Science. Empirical Press, Vancouver, BC (2004) [Электронный ресурс]. Режим доступа: [https://www.researchgate.net/publication/257228469\\_The\\_Development\\_of\\_Social\\_Network\\_Analysis\\_A\\_Study\\_in\\_the\\_Sociology\\_of\\_Science](https://www.researchgate.net/publication/257228469_The_Development_of_Social_Network_Analysis_A_Study_in_the_Sociology_of_Science) Linton C Freeman Empirical Press Vancouver\_BC\_2004 (дата обращения: 23.11.20).



**Беген Петр Николаевич**

Год рождения: 1996  
Университет ИТМО,  
Институт дизайна и урбанистики,  
аспирант группы № 7738,  
направление подготовки: 05.13.10 – Управление  
в социальных и экономических системах,  
e-mail: petyabegen@mail.ru



**Низомутдинов Борис Абдуллохонович**

Год рождения: 1990  
Университет ИТМО,  
Институт дизайна и урбанистики,  
аспирант группы № 7738,  
направление подготовки: 05.13.10 – Управление  
в социальных и экономических системах,  
e-mail: boris@itmo.ru

УДК 004

**РАЗРАБОТКА МЕТОДА ИЗВЛЕЧЕНИЯ АДРЕСА ИЗ ТЕКСТА  
СООБЩЕНИЙ О ПРОИСШЕСТВИЯХ В СОЦИАЛЬНЫХ СЕТЯХ  
ДЛЯ ФОРМИРОВАНИЯ ГЕОПРИВЯЗАННЫХ ДАННЫХ**

**П.Н. Беген, Б.А. Низомутдинов**

**Научный руководитель – к.т.н., директор С.А. Митягин**

Работа выполнена в рамках темы НИР Университета ИТМО №620179 «Разработка картографического сервиса мониторинга потребностей жителей в развитии инфраструктуры городской среды с применением автоматизированных систем обработки данных из социальных сетей».

**Аннотация**

В работе рассмотрена возможность выделения адреса и типа происшествия из текстового массива данных (постов в социальных сетях). Проведен эксперимент с выгрузкой данных из сообщества «ДТП и ЧП | Санкт-Петербург | Питер Онлайн | СПб», всего было собрано и обработано 48 407 записей, далее проведено автоматическое распознавание темы каждого поста и выделение адреса, при его наличии.

**Ключевые слова**

Распознавание текста, машинное обучение, парсер, АПК Безопасный город.

По данным отчета Digital 2020, выпущенного We Are Social и Hootsuite, в 2020 году в Российской Федерации количество интернет-пользователей составило 118 миллионов [1]. Следовательно, 81% граждан Российской Федерации пользуется интернетом.

На январь 2020 года количество пользователей социальных сетей составило 70 миллионов – это 48% от всего населения страны. Количество пользователей социальных сетей в Российской Федерации было неизменно с апреля 2019 года по январь 2020 года.

И теперь интернет – это не только развлечение, многие решают социальные вопросы, получают государственные услуги, отправляют документы в банк и органы

власти [2]. Пользователи генерируют всевозможные данные, обсуждают различные вопросы. Часть из них посвящены происшествиям в городе

На примере Санкт-Петербурга мы рассмотрели возможность извлечения данных о происшествиях, о которых пишут граждане. В качестве площадки для исследования было выбрано сообщество «ДТП и ЧП | Санкт-Петербург | Питер Онлайн | СПб». По состоянию на январь 2021 года, это крупнейшее онлайн сообщество в Санкт-Петербурге о происшествиях, в нем состоит 1 217 175 пользователей, здесь в режиме реального времени пользователи пишут о авариях, угонах и на другие темы.

Цель данной работы - изучение и анализ применения методов извлечения адреса из текстовых сообщений о происшествиях в социальных сетях для формирования геопривязанных данных в формате GeoJSON, которые в дальнейшем можно использовать в ГИС системах или АПК «Безопасный город».

Результатом работы данного подхода может стать распределение происшествий в городе на карте, так как в тексте выделяется 2 сущности: адрес и тип происшествия. На данном этапе выбрано 2 темы - угон и авария. Используя данный подход, мы можем выделить самый аварийный или проблемный участок или найти район, где больше всего происходит угонов. Данная информация есть в официальных источниках, но она не всегда доступна урбанистам и исследователям. Данный подход позволит расширить массив данных для исследователей и граждан.

Кроме того, полученные данные могут быть использованы в АПК «Безопасный город», так как часто сообщение о происшествии поступает не сразу в систему, а данный подход генерирует информацию в режиме онлайн. К примеру, информация о ДТП может попасть в статистику через неделю или даже месяц, если авария была оформлена по европротоколу. Но граждане оперативно фиксируют ДТП и вносят их в данное сообщество, что позволяет искать проблемные места в городе в режиме онлайн.

АПК «Безопасный город» используется для улучшения безопасности среды пребывания людей, а также общественной безопасности, благодаря использованию инновационных информационных систем, которые проводят мониторинг, прогнозируют и, в случае возникновения возможных угроз, предупреждают население. Можно не ограничиваться только такими темами и источниками. Например, в Китае с помощью геоданных из твиттера изучают локальные наводнения [3].

Одной из задач исследования также стала разработка инструментария для автоматической обработки и анализа массива данных постов из социальных сетей. Основной функциональной возможностью инструментария стало автоматическое определение темы поста и распознавание адреса (улица, квартал, дом, район и т.п.). Для получения первичных результатов работы инструментария было принято решение распознавать две темы постов: Угон и Авария. В дальнейшем планируется добавить следующие темы: Пожар, Ограбление, Нападение.

Для разработки инструментария использовалась нейросетевая библиотека *natasha 1.4.0.*, обновленная в 2020 году, для распознавания сущностей в русскоязычном тексте. Для распознавания двух тем «Угон» и «Авария» был настроен *yargy*-парсер. Для извлечения пользовательских сущностей в *yargy* заводятся специальные правила с помощью контекстно-свободных грамматик и специализированных словарей. В рамках исследовательской работы были добавлены простые правила с готовыми предикатами парсера, распознающие слова для выделения темы: «угон, угнали» – Угон, «авария, ДТП, притерли, паравозик, столкнулись, от удара» – Авария.

Для анализа была сформирована выборка данных из постов в сообществе «ДТП и ЧП | Санкт-Петербург | Питер Онлайн | СПб» ([https://vk.com/spb\\_today](https://vk.com/spb_today)). Всего было собрано 48 943 записи.

Для собранной выборки была проведена предварительная обработка постов, которая заключалась в удалении записей, начинающихся со слов «Новости нашего

мегаполиса:». В данном случае мы не учитывали ежедневные сводки новостей при формировании итоговой выборки для анализа. Размер итоговой выборки после удаления таких постов составил 48 407 записей.

Для распознавания тем постов использовались самописные правила для уагу-парсера. В результате была получена выборка данных постов, которые содержали только тему Угон или Авария. На рисунке представлено соотношение постов с двумя выделенными темами к общему количеству записей (рисунок).



Рисунок. Количество постов с темами Угон или Авария

На основе данной выборки была рассчитана средняя точность (ассигасу) распознавания двух выбранных тем, результаты представлены в табл. 1.

Таблица 1

#### Точность распознавания темы в тексте с помощью библиотеки *natasha*

| Тема   | Распознано | Не распознано | Всего записей | Средняя точность |
|--------|------------|---------------|---------------|------------------|
| Угон   | 570        | 124           | 694           | 82,13%           |
| Авария | 9 957      | 1 587         | 11 544        | 86,25%           |

Полученный результат средней точности определения тем в тексте с помощью настроенного уагу-парсера и библиотеки *natasha* оказался хорошим (>80%), однако сам парсер требует доработок, а также увеличения или усложнения правил парсинга необходимых фраз. Также был отмечен средний уровень продолжительности времени работы парсера, на основе чего можно сделать предположение о медлительности работы алгоритма, и с увеличением выборки данных (более 100 тыс. записей) алгоритм парсера будет работать еще медленнее.

Для распознавания адресов в тексте была использована встроенная функция *AddExtractor* из библиотеки *natasha*. Распознавание проводилось на выборке данных, включающих посты с темами Угон и Авария. Всего в выборке выделено 12 238 записей. Чтобы подсчитать среднюю точность распознавания, было задано условие, что если в тексте поста распознана хотя бы одна часть адреса (например, улица, название, номер дома и т.д.), то адрес считается распознанным. Результаты распознавания адресов представлены в табл. 2.

### Точность распознавания адресов с помощью библиотеки *natasha*

| Сущность | Распознано | Не распознано | Всего записей | Средняя точность |
|----------|------------|---------------|---------------|------------------|
| Адрес    | 9 327      | 2 911         | 12 238        | 76,21%           |

Из таблицы видно, что результат средней точности (ассурасу) распознавания адреса в русскоязычном тексте является удовлетворительным для решения задачи (>75%), однако алгоритм и правила распознавания адресов требуют доработок для повышения точности и качества определения распознанных адресов. При распознавании адреса было также отмечено, что наибольший процент точности достигается при определении названия для типа «улица», например в формате «Московская ул.». Однако, если убрать слово-маркер «ул.», то точность распознавания существенно снизится, даже если рядом будут находиться другие маркеры, например «номер дома», «строение», «корпус» и другие части адреса.

В результате анализа полученных результатов задачи распознавания темы поста и выделения адреса в тексте можно сделать вывод о достаточной возможности использования открытых готовых решений, с функционалом гибкой настройки правил, для выполнения задач подобного рода.

В ходе работы была разработана технология по автоматическому распознаванию темы поста и адреса. Рассмотренный эксперимент подтвердил эффективность выбранного набора инструмента *natasha 1.4.0 + yargu-парсер*. Удалось выделять тематику и адреса из текста постов, что должно позволить использовать эти данные в ГИС системах и АПК Безопасный город. В дальнейшей работе планируется увеличение выборки данных за счет использования методов автоматического сбора и извлечения сущностей, повышение точности извлечения и распознавания тем постов и частей адреса, перевод в формат GeoJSON, а также отображение точек происшествий на карте.

### Литература

1. DIGITAL 2020: THE RUSSIAN FEDERATION [Электронный ресурс]. Режим доступа: <https://datareportal.com/reports/digital-2020-russian-federation> (дата обращения: 10.12.2020).
2. Сколько времени россияне проводят в социальных сетях [Электронный ресурс]. Режим доступа: <https://romir.ru/press/adindexru---skolko-vremeni-rossiyane-provodyat-v-socialnyh-setyah> (дата обращения: 12.12.2020).
3. Urban resilience from the lens of social media data: Responses to urban flooding in Nanjing, China // Elsevier Ltd. Bo Wang, Becky P.Y. Looc, Feng Zhene, Guangliang Xie. P.P. 1-13 <https://doi.org/10.1016/j.cities.2020.102884>.



**Низомутдинов Борис Абдуллохонович**

Год рождения: 1990  
Университет ИТМО,  
Институт дизайна и урбанистики,  
аспирант группы № 7738,  
направление подготовки: 05.13.10 – Управление  
в социальных и экономических системах,  
e-mail: boris@itmo.ru



**Углова Анна Борисовна**

Год рождения: 1989  
РГПУ им. А.И. Герцена,  
Институт психологии,  
ассистент кафедры психологии профессиональной деятельности,  
e-mail: anna.uglova@list.ru

УДК 004

**МЕТОД МОНИТОРИНГА ПОТРЕБНОСТЕЙ  
ЖИТЕЛЕЙ В РАЗВИТИИ ГОРОДСКОЙ ИНФРАСТРУКТУРЫ**

**Б.А. Низомутдинов, А.Б. Углова**

**Научный руководитель – к.т.н., директор С.А. Митягин**

Работа выполнена в рамках темы НИР Университета ИТМО № 620179 «Разработка картографического сервиса мониторинга потребностей жителей в развитии инфраструктуры городской среды с применением автоматизированных систем обработки данных из социальных сетей».

**Аннотация**

В работе рассмотрена возможность использования данных из социальных сетей для ценностно-ориентированного управления программами развития города. Рассматривается теоретическое и экспериментальное изучение ценностно-смысловых аспектов городского пространства Санкт-Петербурга на основе автоматизированной выгрузки данных из социальных сетей, для мониторинга потребностей жителей в развитии инфраструктуры городской среды на разных уровнях централизации.

**Ключевые слова**

Управление городом, парсинг, социальные сети, город.

Для обеспечения устойчивого развития и высокого качества жизни в городской среде, необходимо применения интеллектуальных городских технологий и особое внимание к практике учета эко-психологических технологий [1, 2]. Одним из самых эффективных подходов управления развития инфраструктуры города можно считать ценностно-ориентированный подход [3].

Развитие крупного города как динамичный процесс нуждается в постоянном контроле и расширении возможностей коммуникации человека и среды для

мониторинга потребностей жителей, а также оперативного реагирования и решения социальных и экономических проблем.

Часто система управления городом не отвечает требованиям современности, управление происходит на плановой основе или используются устаревшие статистические данные, отсутствует четкий механизм изменения стратегии в результате изменения внешней среды. Данные проблемы, как правило, возникают из-за недостатка информации или отсутствия у органов власти обратной связи от граждан.

В исследовании рассматривается возможность использования данных из социальных сетей в качестве источника информации с целью выявления актуальных запросов граждан в развитии городской среды.

Как следует из данных сборника «Цифровая экономика: 2020» [4], выпущенного Институтом статистических исследований и экономики знаний Высшей школы экономики в партнерстве с Минкомсвязью России и Росстатом, ежедневно в интернет заходят 68,8% россиян в возрасте 15–74 лет, в том числе 88,9% молодежи 15–24 лет.

И теперь, времяпрепровождение в интернете – это не только развлечение, многие решают социальные вопросы, получают государственные услуги, отправляют документы в банк и органы власти. Социальные сети, отличный канал информационного взаимодействия граждан и органов власти. Современный житель мегаполиса большую часть информации о своем районе, культурных, социальных и политических мероприятиях узнает из социальных сетей, что делает данные сетевые коммуникационные площадки наиболее информативными для экспертного изучения [5].

Для разработки метода изучения социально-культурных и экономических потребностей жителей разных районов города нами был подготовлен следующий план работ:

- проанализировать основные подходы к изучению семиотики городской среды, и определить методологию исследования информационного образа города на основе информации, представленной в социальных сетях.
- сформировать репрезентативный корпус городских сообществ, представленных в социальных сетях, на примере сообществ разных районов Санкт-Петербурга.
- настроить комплекс автоматизированного сбора данных об участниках виртуальной коммуникации в городских сообществах.
- выявить структуру и содержание городских сообществ на примере сообществ разных районов Санкт-Петербурга.
- провести психодиагностическое обследование испытуемых, состоящих в различных городских сообществах, для определения основных интересов и потребностей, опосредующих взаимодействие с городской средой.
- провести автоматизированный анализ профилей испытуемых в социальных сетях, для выявления основных ценностно-смысловых компонент их информационного образа, являющегося основой для построения коммуникации в городском пространстве
- теоретическое и экспериментальное изучение ценностно-смысловых аспектов городского пространства Санкт-Петербурга.
- разработать картографический сервис для мониторинга потребностей жителей в развитии инфраструктуры городской среды на разных уровнях централизации, на основе экспериментального изучения их ценностно-смысловой сферы (публичный сервис – вебсайт).

На первом этапе исследования мы провели изучение цифрового образа. Было отобрано 300 респондентов, им было предложено пройти психодиагностическое тестирование. После этого, испытуемые указывали свой адрес в социальной сети Вконтакте. Далее автоматизированным способом была выгружена вся доступная

информации из профиля, и с помощью корреляционного анализа были выявлены множественные взаимосвязи с психологическими характеристиками пользователей, а также с личностными свойствами респондентов. На основе полученных взаимосвязей был сделан вывод о приемлемых возможностях автоматизированного анализа цифровой социальной идентичности, а также о целесообразности в продолжении работ.

На следующем этапе исследования нами было отобрано 500 городских сообществ во всех районах города и выгружена вся информация о подписчиках, представленная в виде ID адресов, база данных составила более 4 миллионов записей.

На рисунке представлена общая модель сбора данных, которая проведена в ходе исследования.

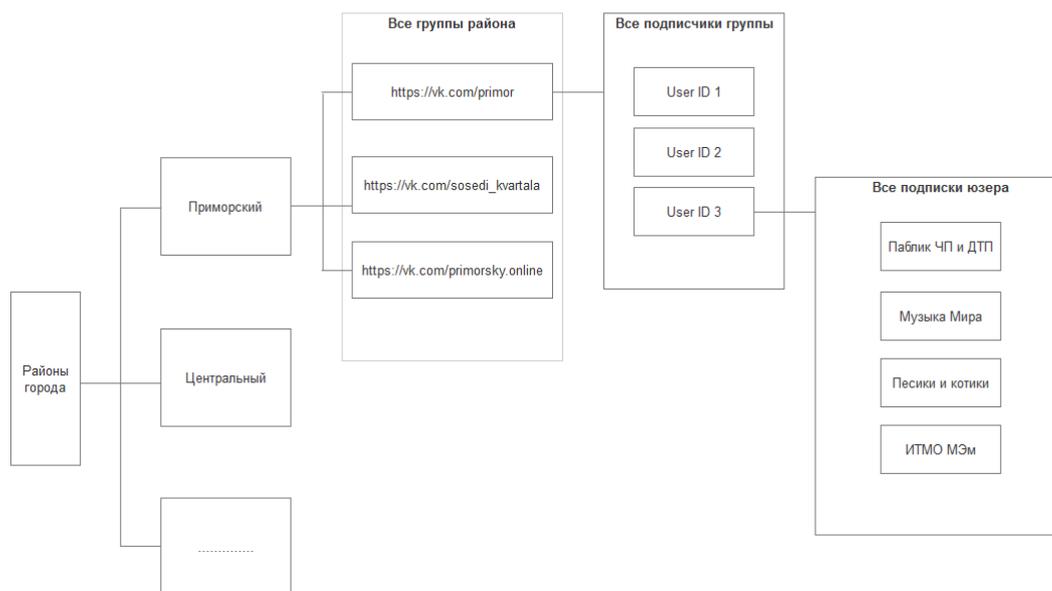


Рисунок. Общая схема сбора информации

Далее запланирована выгрузка основной информации профилей. Это позволит получить основные ценностно-смысловые компоненты информационного образа, а также базовые социально-демографические характеристики. Учитывая, что мы знаем, из какого района конкретный профиль, мы можем осуществить распределение результатов по районам и выгружать данные в ГИС систему для визуальной аналитики.

Важным компонентом системы станет модель деперсонализации собранных данных. В ходе исследования мы не храним фамилию и имя подписчиков, а используем только статистическую информацию.

На следующем этапе исследования запланировано провести психодиагностическое обследование испытуемых из нашей базы, отнесенных к разным районам города, для определения основных интересов и потребностей, опосредующих взаимодействие с городской средой, далее будет проведен автоматизированный анализ профилей испытуемых в социальных сетях для выявления основных ценностно-смысловых компонент их информационного образа, являющегося основой для построения коммуникации в городском пространстве.

Используя полученные данные, запланировано разработать картографический сервис программы мониторинга потребностей жителей в развитии инфраструктуры городской среды на разных уровнях централизации.

Применение новых технологий и методов прогнозирования позволит повысить точность анализа изменений внутренней среды городского пространства, прогнозирования нужд и потребностей жителей города, а также даст способность системам управления корректировать задачи и оперативные планы изменения инфраструктуры под новые обстоятельства.

Подобные технологии управления активно внедряются в международной практике. Создание и использование предсказательных моделей на основе социальных сетей является одним из самых перспективных направлений создания и тестирования новых городских технологий. Во всем мире популярным становится широкое использование и внедрение в различные сферы городского хозяйства информационных технологий для решения городских проблем и управления, что получило название «Умный город» или «Smart City». Данное исследование поможет в развитии систем Умного города, направленных на оперативный мониторинг актуальных потребностей населения.

Создание картографического сервиса позволит приблизиться к решению одной из интереснейших проблем социально-психологического знания – выявления взаимосвязи социальной идентичности, информационной среды и реальной среды обитания пользователей в реальном времени, что может создать предпосылки для совершенствования успешной коммуникации человека и города.

### Литература

1. Yigitcanlar T. Smart cities: an effective urban development and management model? // *Australian Planner* 52(1) PP 27-34, 2015, DOI: 10.1080/07293682.2015.1019752.
2. Van Oers, R.; Pereira Roders, A. Historic cities as model of sustainability. *J. Cult. Herit. Manag. Sustain. Dev.* 2012, 2, PP 4–14.
3. Hurlings L.G. Values in place; A value-oriented approach toward sustainable place-shaping, 2014, PP. 257-274.
4. Цифровая экономика: краткий статистический сборник // [Электронный ресурс]. Режим доступа: <https://issek.hse.ru/mirror/pubs/share/323871553> (дата обращения: 25.12.2020).
5. Королева Н.Н., Богдановская И.М., Углова А.Б. Психосемиотика повседневного образа города в автобиографических нарративах петербуржцев // *Известия Российского государственного педагогического университета им. А.И. Герцена.* 2018. № 187. С. 177-188.



**Николаева Ксения Алексеевна**

Год рождения: 1994  
Университет ИТМО,  
факультет цифровых трансформаций,  
студент группы №С42131,  
направление подготовки: 09.04.02. – Цифровое здравоохранение,  
e-mail: k.nikolaeva@itmo.ru



**Ковальчук Сергей Валерьевич**

Год рождения: 1984  
Университет ИТМО,  
факультет цифровых трансформаций,  
к.т.н.,  
e-mail: kovalchuk@itmo.ru

УДК 519.876.5

## **МОДЕЛИРОВАНИЕ ИНФОРМАЦИОННЫХ ПРОЦЕССОВ В СЛОЖНЫХ РАСПРЕДЕЛЕННЫХ ТЕЛЕМЕДИЦИНСКИХ СИСТЕМАХ**

**К.А. Николаева**

**Научный руководитель – к.т.н. С.В. Ковальчук**

Работа выполнена в рамках темы НИР №620171 «Интеллектуальные технологии моделирования поведения агентов системы здравоохранения».

### **Аннотация**

В работе описывается создание базовой дискретно-событийной модели информационной медицинской системы для обработки показателей давления пациентов с артериальной гипертензией, создание правил обработки показателей и оповещения оператора системы, что является необходимым для задачи определения валидности событий и дальнейшей работы с пациентами.

### **Ключевые слова**

МИС, моделирование процессов, телемедицина, дискретные модели, SimPy.

В настоящее время телемедицина как никогда актуальна и востребована: постоянный рост населения, увеличение демографического старения и многие другие причины создают огромную нагрузку на систему здравоохранения. Несмотря на то, что современные технологии снижают нагрузку на медицинский персонал и автоматизируют многие рутинные процессы, ученые сталкиваются с новой проблемой: при постоянном масштабировании система становится сложной в управлении и затрудняется обработка данных в ней, поэтому ставится вопрос о необходимости создания децентрализованных информационных систем [1].

Для создания распределенной системы требуется предсказать информационные потоки, слабые места системы, какие моменты возможно контролировать в

автоматическом режиме, чтобы снизить нагрузку на реальных операторов медицинских систем [2].

В данной работе была реализована базовая дискретно-событийная модель информационного взаимодействия с оценкой информации о пациентах, полученной в процессе моделирования. Схема модели указана на рис. 1, где отмечена базовая модель и дальнейшая работа по расширению ее.



Рис. 1. Схема информационной модели

Модель реализована на языке программирования Python с использованием библиотеки имитационного моделирования SimPy. В модели реализованы три основные блока функций:

1. Расписание. Блок генерирует расписание, по которому запрашивается давление от пациентов при моделировании.

2. Генерация давления. По запросу с помощью базового класса `rv_continuous` и метода `rvs` (случайные переменные) создается систолическое значение давления пациента в зависимости от принадлежности к группе.

3. Обработка значений. Блок обрабатывает полученные значения давления в ходе выполнения сценария моделирования и выдает сообщения при выполнении заданных базовых правил.

На рис. 2 отражена блок-схема реализации сценария моделирования.

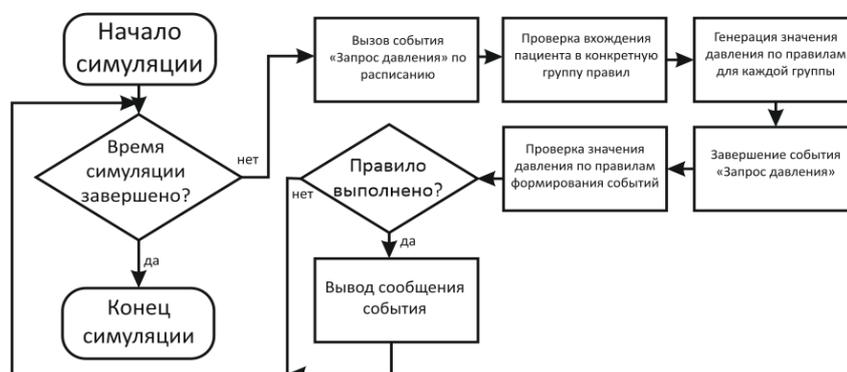


Рис. 2. Блок-схема выполнения моделирования

Сценарий симуляции включает в себя две группы пациентов: «здоровые» и «больные» по 5 человек в группе; расписание замеров: 2 раза в день; период моделирования: 30 дней.

В ходе выполнения сценария выводились сообщения системы об ошибочных измерениях (рис. 3) и получен график динамики давления пациента с полученными ошибочными значениями (рис. 4).

```
id: 4 . day 4 : error measurement  
id: 3 . day 5 : error measurement  
id: 7 . day 5 : error measurement  
id: 1 . day 6 : error measurement  
id: 5 . day 6 : error measurement  
id: 0 . day 7 : error measurement  
id: 5 . day 7 : error measurement  
id: 3 . day 8 : error measurement  
id: 9 . day 8 : error measurement  
id: 2 . day 9 : error measurement  
id: 3 . day 9 : error measurement  
id: 8 . day 9 : error measurement
```

Рис. 3. Фрагмент сообщений системы об ошибках

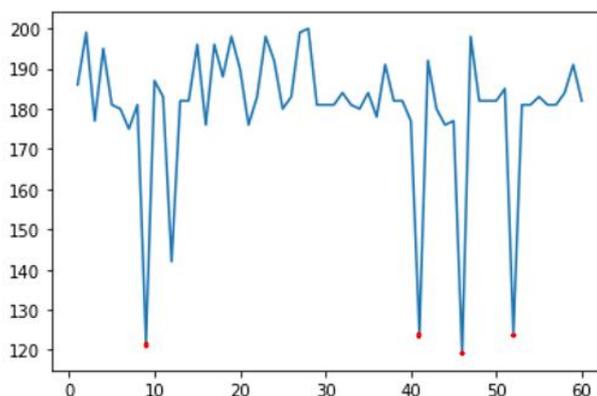


Рис. 4. График давления пациента id 8

Из результатов видно, что базовая модель справляется с задачей определения ошибочных показателей давления, однако стоит отметить, что базовые правила обработки весьма грубы и на реальных значениях давления могут обрабатывать события ошибочно. Так, на примере графика динамики давления (рис. 4) видно, что существует выброс, не отмеченный системой, однако в реальной жизни может быть существенным при определении состояния пациента.

В дальнейшей работе генерация давления будет зависеть от набора параметров пациента (пол, возраст, вес), а также планируется реализация более сложных правил обработки значений давления.

### Литература

1. Xie T. et al. A system dynamics approach to understanding the One Health concept // PLoS One. 2017. Vol. 12, № 9. P. 1–11.
2. Abuhay T.M. et al. Simulation, analysis, and prediction of medical care processes for cardiological in-patients // Izv. vysshih uchebnyh Zaved. Priborostr. 2018. P. 730–733.



**Огирь Константин Сергеевич**

Год рождения: 1998  
Университет ИТМО,  
факультет инфокоммуникационных технологий,  
студент группы №К41103с,  
направление подготовки: 11.04.02 – Программное  
обеспечение в инфокоммуникациях,  
e-mail: Kostya\_ogir@mail.ru



**Осипов Никита Алексеевич**

Год рождения: 1972  
Университет ИТМО,  
факультет инфокоммуникационных технологий,  
к.т.н., доцент,  
e-mail: nikita@ifmo.spb.ru

УДК 004.8

**ПЕРСПЕКТИВЫ ИСПОЛЬЗОВАНИЯ МАШИННОГО  
ОБУЧЕНИЯ ПРИ АНАЛИЗЕ ПРОГРАММНОГО КОДА**

**К.С. Огирь**

**Научный руководитель – к.т.н., доцент Н.А. Осипов**

Работа выполнена в рамках темы НИР «Исследование способов применения методов машинного обучения для оценки качества кода».

**Аннотация**

В работе рассмотрена возможность применения машинного обучения в анализе программного кода. Определены основные проблемы применения машинного обучения в анализе программного кода, возможные пути решения найденных проблем и будущие перспективы его применения.

**Ключевые слова**

Машинное обучение, анализ программного кода, статические анализаторы, проблемы применения машинного обучения, перспективы применения машинного обучения.

С ростом сложности и объема написанного программного кода становится всё сложнее уследить за всеми аспектами разрабатываемого программного продукта (ПП). И внесение изменений или обнаружение дефектов в ПП на поздних этапах разработки может потребовать от разработчиков больших временных и трудовых вложений, что влечет за собой удорожание производства такого продукта. Поэтому для избегания подобных дополнительных трат при производстве сейчас при написании программного кода программисты и архитекторы программ обращают огромное внимание на то, как написан код, т.е. учтены ли всевозможные будущие изменения, написана ли документация к каждому методу, заканчивая тем, что разработчики думают, как кратко и ёмко дать название той или иной переменной и методу, где стоит выравнять написанный код, а где оставить пустую разделяющую строку. Но и это не решает проблемы позднего обнаружения дефектов программного кода, для этого разработчиками производится анализ кода – код-ревью (code review), который является

эффективным, но трудозатратным и небыстрым процессом. Однако для анализа написанного кода можно использовать также программные анализаторы.

Существующее множество различных анализаторов исходного программного кода можно разделить на 2 группы: статические и динамические анализаторы. Статические анализаторы позволяют проанализировать код без его непосредственного выполнения, рассматривая непосредственно сам программный код и файлы метаданных [1]. Динамические анализаторы, в свою очередь, анализируют программу посредством ее выполнения и отслеживания поведения программы в различных ситуациях [2].

Некоторые аспекты статического анализа, такие как расчет метрик, обратный инжинерия и др., достаточно просты в реализации и использовании, поэтому зачастую встраиваются в среды разработки, такие как Visual Studio, Intel IDEA, Eclipse и др. Более сложные формы статического анализа могут представлять отдельные программы, расширения и подключаемые программные пакеты. Как следствие такой простоты использования и доступности, статические анализаторы приобрели наибольшую популярность, т.к. позволяют обнаруживать и устранять ошибки и уязвимости на самых ранних этапах – в ходе разработки. Наиболее популярными алгоритмами статического анализа кода являются: интервальный анализ; анализ указателей; ресурсный анализ; строковый анализ; анализ зависимостей.

Вместе с этим статические анализаторы накладывают свои ограничения на их использование: для анализа программы статическим анализаторам необходим доступ ко всему анализируемому коду для более корректной его оценки; результаты, выдаваемые каким-либо анализатором, не являются на 100% верными, т.к. могут давать ложноположительные предупреждения, которые должны рассматриваться разработчиками, и именно ими должно приниматься решение о дальнейших действиях; когда речь идет о крупных проектах, с большой кодовой базой невозможно провести за адекватное время полный всесторонний анализ системы, что влечет за собой ограничение области анализа или его точности.

В тех случаях, где статические анализаторы требуют больших ресурсов для анализа кода, динамические анализаторы справляются куда лучше, например, выявление утечек памяти и параллельных ошибок. При динамическом анализе проверяют всевозможные пути выполнения программы путем генерации всевозможных входных данных. На основе построенной карты маршрутов при статическом и динамическом анализе программы проверяются функциональные требования. При этом динамический анализ не предусмотрен на этапе разработки, а большое количество сочетаний входных данных программы ведет к экспоненциальному росту их объема, что требует от подобных анализаторов возможности параллельного запуска потоков выполнения программы с различными входными данными, возможности анализировать только определённые пути.

Машинное обучение в анализе программного кода может применяться для решения различных задач, начиная от поиска дефектов кода до автоматической генерации тестов на основе уже имеющихся и поиска «запахов кода» [3]. Основным смыслом применения машинного обучения в анализе кода – классификация кода или участков кода по определенным параметрам, т.е. по набору параметров кода (предикторов). Разберем подробнее возможности применения машинного обучения: поиск дефектов (уязвимостей и ошибок) в программном коде, поиск повторяющегося кода, анализ программ, поиск шаблонов (паттернов) в программном коде, рефакторинг [4]. Также при помощи машинного обучения осуществляется генерация документации, трансляция программного кода на языки программирования другого типа, генерация программного кода на основе спецификаций, а также генерация описания на

естественном, понятном человеку языке того, что делает программный код, и даже генерация рекомендаций по написанию кода.

Поиск дефектов в программном коде – поиск распространенных ошибок в программном коде, и является одним из наиболее явных применений машинного обучения в анализе кода. При этом рассматриваются различные модели применения машинного обучения: нейронные сети, случайный лес, деревья принятия решений, байесовские сети доверия [5]. Поиск шаблонов и запахов в программном коде позволяет выявлять определенные шаблоны в программном коде, понятные человеку, указывающие на определенные проблемы в программном коде, которые в дальнейшем усложнят разработку или поддержку программного продукта. Для идентификации различных запахов и паттернов проектирования используют методы случайного леса, наивного байесовского классификатора, дерева решений и методов опорных векторов.

Важным аспектом при использовании машинного обучения являются те наборы данных, на которых производится обучение модели машинного обучения. От качества и количества примеров в наборе данных напрямую зависит и результат работы модели. Существуют уже готовые кодовые базы для обучения моделей: BegBunch, OpenSolaris, открытая кодовая база NASA и др. Проверенные наборы данных программного кода из реальных проектов или специально составленные являются эталонными и предоставляют большое количество вариаций одного и того же дефекта или запаха кода, поэтому большинство исследователей используют именно доступные проверенные источники в своих работах. Важно отметить, что распространение open-source проектов, т.е. проектов с открытым исходным кодом, над которыми постоянно происходят улучшения, позволяет получать реальный рабочий код программ с комментариями разработчиков. Для обучения модели машинного обучения поиска различных запахов кода, дефектов необходимо заранее определить предикторы – изменяемые параметры в машинном обучении, на основании которых определяется результат работы модели машинного обучения. Такими предикторами для поиска запахов кода могут быть как обычные статистические метрики кода, описанные в предыдущих разделах, так и результаты, полученные при помощи статических анализаторов.

Для решения проблемы переобучения модели машинного обучения в анализе программного кода обычно используется метод перекрестной проверки, суть которой заключается в разделении исходной обучающей выборки на  $n$  частей, после чего производится обучение на  $n-1$  части и тестирование модели на оставшейся. Алгоритм проводит обучение и проверку для каждой из  $n$  частей. За счет использования такого метода достигается равномерное использование данных, что решает проблему переобучения модели. Вместе с тем такой алгоритм позволит более эффективно использовать имеющиеся данные, если примеров кода будет не хватать.

Несмотря на большое количество исследований и экспериментальных анализаторов, большинство исследователей склоняются к тому, что полная замена существующих алгоритмов анализа программ методами машинного обучения невозможна по многим причинам. Во-первых, – проводя анализ, было замечено что в каждом из них рассматриваются сразу несколько или несколько десятков различных способов классификации. Это объясняется тем, что нет доказанных утверждений о превосходстве одного метода над другими, выделяют несколько методов, результаты которых лучше, чем у других, однако нет никаких четких обоснований отказываться от использования остальных. Во-вторых, в основе машинного обучения стоят предикторы – метрики исходного кода, определяющие результат модели, и суммарное количество которых может превышать сотню. Так же, как и для различных методов машинного обучения, нет общих соглашений о наиболее результативной комбинации предикторов для каждой решаемой задачи. В-третьих, как показывают исследования, и реальная

практика результат работы анализатора с применением МО зависит от той кодовой базы, на которой обучался алгоритм, лежащий в основе этого анализатора. Исследователями было замечено, что модель МО, обученная на одной обучающей выборке, хуже справляется с примерами из тестовой выборки, не включенной в обучающую. А также обученный на искусственных примерах (примерах кода, составленных специально для обучения, а не взятых из реальных проектов) плохо справляется с примерами из реальных рабочих проектов и наоборот. В-четвертых, необходимо формировать обучающую выборку для каждой классификации, выполнение этой задачи вручную займет слишком большое количество ресурсов.

Применение машинного обучения в анализе кода показывает нестабильные результаты по причине обилия проблем, указанных ранее. Поэтому на текущий момент машинное обучение имеет перспективы дополнительной меры контроля в статических анализаторах, т.к. статистика показывает, что количество истинно-положительных результатов при применении методов машинного обучения может достигать 95%. Также методы машинного обучения слабо подвержены выбросам и легко справляются с шумом в данных, что можно рассматривать как положительную сторону МО при анализе программного кода, в котором одна и та же функция может быть написана множеством различных способов, используя разные операторы, команды которой могут располагаться в различных последовательностях. С подсчетом метрик и их анализом хорошо справляются существующие методы статического анализа, однако для обработки сложных структур и нахождения зависимостей в них статическим анализаторам требуются сложные алгоритмы и большое количество ресурсов. Однако для решения этой проблемы как нельзя лучше подходят методы машинного обучения, поэтому одной из перспективных направлений машинного обучения в анализе кода является синтаксическая обработка кода, т.е. выявление паттернов и анти-паттернов, генерация кода и документации на основе кода, рефакторинг – обработка программного кода как естественного языка и выявление в нем определенных зависимостей. Машинное обучение в целом уже косвенно влияет на написание программного кода, например, посредством рекомендательных систем для разработчиков, предлагая автозавершение кода, т.к. многие из таких систем используют методы интеллектуального анализа и машинного обучения. Другим примером могут служить системы генерации тестов на основе кода и уже имеющихся тестов, позволяя более полно покрывать программный код тестами с различными входными параметрами и результатами. Согласно исследованиям, методы машинного обучения уже могут применяться для узкого круга дефектов кода, а также для поиска повторяющихся элементов исходном коде. Также рассматриваются нестандартные подходы к анализу программного кода – т.к. лексика языков программирования основана на лексике естественного языка, следовательно, можно рассматривать программный код как набор упорядоченных слов и сокращений. И ввиду того, что нейронные сети хорошо себя показывают при обработке естественного языка, применение нейронных сетей и глубокого обучения могут дать хорошие результаты в анализе программного кода.

В ходе работы была рассмотрена возможность применения методов машинного обучения для анализа программного кода. В результате выполнения работы были выделены основные проблемы, мешающие применению машинного обучения в анализе кода. А также были определены перспективы дальнейшего применения машинного обучения в анализаторах программного кода.

Целью дальнейших исследований будет решение выявленных проблем, а также исследование и разработка перспективных направлений применения машинного обучения в анализе программного кода.

### Литература

1. Barstad V., Goodwin M., Gjørseter T. Predicting Source Code Quality with Static Analysis and Machine Learning // Norsk IKT-konferanse for forskning og utdanning: материалы конф. Fredrikstad. 2014. С. 12.
2. Лебедев В.И. Методы и перспективы статического анализа программного кода // Молодежный научно-технический вестник. 2015. №12. С. 28.
3. Detecting code smells using machine learning techniques: Are we there yet? / Dario Di [и др.] // IEEE 25th International Conference on Software Analysis, Evolution and Reengineering (SANER). 2018. С. 612-621.
4. A Survey of Machine Learning for Big Code and Naturalness / Miltiadis A. [и др.] // ACM Computing Surveys. 2018. Т.51. С. 1-37.
5. Machine learning for finding bugs: An initial report / Gevay S. [и др.] // 2017 IEEE Workshop on Machine Learning Techniques for Software Quality Evaluation (MaLTeSQuE). 2017. С. 21-26.



**Панфилов Георгий Олегович**

Год рождения: 1995

Университет ИТМО,

Институт дизайна и урбанистики,

студент группы №С42042

направление подготовки: 09.04.03 – Прикладная информатика,

e-mail: panfilovgeorg@mail.ru

УДК 004

## **СРАВНИТЕЛЬНЫЙ АНАЛИЗ ПРОГРАММНОГО ОБЕСПЕЧЕНИЯ ДЛЯ АВТОМАТИЗИРОВАННОГО СБОРА ИНФОРМАЦИИ ИЗ СОЦИАЛЬНЫХ СЕТЕЙ**

**Г.О. Панфилов**

**Научный руководитель – к.полит.н. А.В. Чугунов**

Работа выполнена в рамках НИР Университета ИТМО №620179 «Разработка картографического сервиса мониторинга потребностей жителей в развитии инфраструктуры городской среды с применением автоматизированных систем обработки данных из социальных сетей».

### **Аннотация**

Статья посвящена анализу программного обеспечения для автоматизированного сбора информации из социальных сетей. Описывается область применения данного программного обеспечения в рамках исследовательских задач, связанных с обработкой жалоб граждан в официальных группах органов власти в социальных сетях. Далее анализируются ключевые функциональные возможности подобных программ, предлагается их классификация и описывается методика оценивания. Также производится сравнительный анализ парсеров каждого из описанных типов. В заключение делаются выводы об актуальности их применения.

### **Ключевые слова**

Социальная сеть, автоматизированный сбор данных, парсер.

Социальные сети сегодня являются одним из основных носителей открытой информации и активно используются для анализа социального поведения. Как следует из данных сборника «Цифровая экономика: 2020», выпущенного Институтом статистических исследований и экономики знаний Высшей школы экономики в партнерстве с Минкомсвязью России и Росстатом, ежедневно в Интернет заходят 68,8% россиян в возрасте 15–74 лет, в том числе 88,9% молодежи 15–24 лет. Наиболее востребованный вариант использования - пребывание в социальных сетях. По данным исследовательского холдинга «Ромир», среднестатистический пользователь находится в социальных сетях в среднем 143 минуты в день.

Автоматизированный сбор данных из глобальной сети Интернет, в том числе из социальных сетей, сегодня становится все более и более востребованным. Быстрый сбор большого объема данных актуален с целью решения различных задач – от подбора целевой аудитории и мониторинга общественного мнения до вопросов безопасности и предотвращения правонарушений [1, 2]. Данные социальных сетей активно анализируются исследователями для изучения городского пространства [3, 4].

Существует несколько основных методов автоматизированного сбора информации из социальных сетей. Два наиболее распространенных метода – это:

1. Получение необходимых данных через интерфейс программирования приложений (API – Application Programming Interface). API – проблемно-ориентированный интерфейс взаимодействия между сайтом и сторонними программами.

2. Выгрузка части html-кода по заранее заданным признакам посредством использования программы выгрузки (парсинг).

Каждый из способов имеет свои преимущества и недостатки. Основными недостатками применения API является: его возможное отсутствие в исследуемом сервисе. Также в ряде случаев это может быть неоправданно дорого. Оба способа могут быть применены в зависимости от целей и задач исследования. В нашем случае представляется целесообразным использовать второй способ.

В сущности своей парсинг представляет собой синтаксический анализ информации, опубликованной в сети. Содержимое интернет-страниц является иерархичным набором данных, структурированным с помощью человеческих и компьютерных языков. Контент большинства веб-страниц создается при помощи языка HTML, который позволяет отображать документы в удобной для человека форме. Парсинг веб-страниц применяется для получения в чистом виде текстовой информации из статей, размещенных на страницах сайтов [5].

Линейка программного обеспечения предназначенного для парсинга контента довольно обширна. Программы фокусируются на сборе разных данных и имеют разный функционал, но суть их сводится к одному – к ускоренному копированию информации на однотипных страницах с выгрузкой части HTML-кода по заранее заданным признакам.

Задача проведения сравнительного анализа программного обеспечения для автоматизированного сбора информации из социальных сетей возникла в исследовательской группе Центра технологий электронного правительства Института дизайна и урбанистики Университета ИТМО. В процессе работы над проектом, связанным с изучением комментариев граждан в официальных группах органов власти в социальных сетях, потребовался инструмент для автоматизированного сбора комментариев и проведения анализа профилей пользователей, который бы позволил осуществлять выгрузки данных из социальных сетей по заранее заданным признакам.

После первичного мониторинга рынка программного обеспечения для парсинга была предложена следующая классификация.

Во-первых, были выделены в отдельную группу программы, сфокусированные на парсинге аудитории социальных сетей – например, сборе данных об активных пользователях, сообществах в социальных сетях и изучении целевой аудитории.

Во вторую группу были выделены программы мониторинга ключевых слов, позволяющих решить задачи изучения упоминания интересующего явления в сети.

В третьих, в классификацию были включены универсальные парсеры контента, предназначенные для сбора любого контента, вне зависимости от принадлежности к социальным сетям. Такое программное обеспечение может быть использовано для решения задачи сбора данных с определенного Интернет-ресурса с определенной периодичностью (например, интернет-магазина, форума и т.п.)

Таким образом, для оценки релевантности выбора программы одной из категорий для решения задач исследования потребовалось провести сравнительный анализ нескольких наиболее широко распространенных парсеров каждой категории.

В качестве критериев оценивания было выбрано четыре базовых критерия, одинаково применимые ко всем трем категориям:

1. Стоимость.

2. Функциональные возможности парсинга.
3. Возможности выгрузки (доступные форматы).
4. Основная применимость (для каких целей).

Далее был осуществлен сбор данных о программном обеспечении каждой из категорий.

Результаты сбора данных о программах для парсинга аудитории представлены в табл. 1. Как мы можем видеть, имеются примеры недорогого программного обеспечения (VK.WATCH, PEPPER.NINJA), но предназначенные для работы с одной социальной сетью («ВКонтакте»). Имеются также более сложные инструменты с более высокой стоимостью (VK.BARKOV.NET), в том числе с услугой индивидуальной настройки (CLICK.RU).

Таблица 1

### Сравнение ПО для парсинга аудитории социальных сетей

|                          | VK.WATCH   | PEPPER.NINJA   | VK.BARKOV.NET  | CLICK.RU   |
|--------------------------|--|--|--|--|
| 1. Стоимость             | Полная версия – от 4 евро  | От 490 рублей  | 24 часа – 249 р.<br>1 месяц – 699 р.<br>3 месяца – 1699 р.<br>1 год – 5999 р.                    | Индивидуально  |
| 2. Возможности парсинга  | Работа с историческими данными в «ВК»<br>Самый большой архив исторических данных сервиса VK: комментарии пользователей, фотографии, списки друзей и изменения профилей.<br>Пользователя можно найти по уникальному ID, либо по последнему никнейму | Парсинг аудитории «ВК» с помощью продвинутых алгоритмов. Сбор пользователей по нескольким десяткам признаков: возраст, пол, геоположение, семейное положение, место учебы и работы, проявленная активность, интересы и многое другое | Набор инструментов для сбора самых разнообразных данных из «ВК» и «Одноклассники» в удобном виде | Инструмент создан для сбора активной целевой аудитории в социальной сети «ВК» для запуска таргетированной рекламы. Сбор узкой аудитории с общими интересами. Для сбора пользователей укажите группы по интересам и получите готовый список, который можно загрузить в рекламный кабинет «ВКонтакте» или MyTarget |
| 3. Возможности выгрузки  | CSV  | CSV/TXT/XLS  | CSV/TXT/XLS  | Загрузка списком или XLSX-файлом. Облачное хранение отчетов с постоянным доступом  |
| 4. Основная применимость | «ВКонтакте»  | «ВКонтакте»  | Социальные сети  | «ВКонтакте»  |

Далее были собраны данные о программах, предназначенных для парсинга ключевых слов и анализа частоты упоминания представлены в табл. 2. Также как и в

предыдущем случае, продукты очень разнообразны: имеются как относительно недорогие аналоги (StarComment, Babkee), так и дорогие платформенные решения с более широкими возможностями, предназначенные для крупного бизнеса (YouScan).

Таблица 2

**Сравнение ПО мониторинга упоминаний и ключевых слов в сети**

|                          | StarComment  | YouScan   | Babkee   |
|--------------------------|--|---|--|
| 1. Стоимость             | От 290 до 2000 рублей  | Полный доступ стоит от 35 000 руб. в месяц. Есть бесплатная демонстрация. Похожие решения: IQBuzz                             | Есть бесплатный тариф, у которого с ограничениями: до 2 объектов мониторинга, до 3 000 сообщений по каждому из них |
| 2. Возможности парсинга  | - Мониторинг комментариев «ВК»;<br>- Мониторинг записей «ВК»;<br>- Отслеживание комментариев по ключевым словам (до 100 слов в задании);<br>- Большой выбор объектов отслеживания (профили, сообщества, группы). | Платформа для маркетинговых исследований, технической поддержки, управления репутацией. Подходит больше для крупного бизнеса. | Система мониторинга соцмедиа и средств массовой информации от агентства Ingate. В ее базе — миллионы площадок.     |
| 3. Возможности выгрузки  | TXT, CSV   | TXT, CSV, Excel   | TXT, CSV   |
| 4. Основная применимость | «ВК», «Instagram», «Facebook», «Одноклассники», «YouTube» и др.  | «ВК», «Одноклассники», «Instagram», «Twitter», СМИ и др.  | LiveJournal, Одноклассники, Вконтакте и др.  |

На третьем этапе был проведен сравнительный анализ для универсальных парсеров контента, подходящих как для парсинга социальных сетей, так и для ресурсов иных типов. Были собраны данные о таких продуктах, как Datacol, Content Downloader X1, Octoparse и Parsehub. Результаты представлены в табл. 3. Стоимость программ данной категории значительно выше – стоимость лицензии на продолжительный период составляет от 2-5 тыс. руб. В некоторых решениях (например, Datacol) присутствуют наборы настроек для парсинга разных типов ресурсов, в том числе, социальных сетей.

Таким образом, в рамках работы был произведен сравнительный анализ программного обеспечения трех основных типов (парсеры аудитории, парсеры ключевых слов, универсальные парсеры). Окончательный выбор программы будет осуществлен на следующем этапе исследования в зависимости от бюджета проекта и соответствия программ задачам эмпирической части исследования.

Сравнение уникальных парсеров контента

|                          | Datacol  | Content Downloader X1   | Octoparse   | Parsehub  |
|--------------------------|--|---|---|---|
| 1. Стоимость             | 1 неделя – 795 р.<br>3 месяца 3.4 т.р.<br>6 месяцев – 5 т.р.<br>12 месяцев – 7.5 т.р.                      | Демо-версия –<br>бесплатно<br>Полная версия –<br>2-5 т.р.   | Бесплатно – для<br>небольших<br>проектов<br>(10 тыс. записей<br>на экспорт<br>2 одновременных<br>локальных<br>прогона)<br>58\$/мес. – для<br>небольших<br>команд<br>(Неограниченный<br>экспорт данных<br>Неограниченный<br>одновременный) | Демо-версия -<br>бесплатно:<br>200 страниц<br>данных за 40 мин.<br>5 проектов,<br>Ограниченная<br>поддержка,<br>Хранение данных<br>14 дней<br>Полная версия -<br>149\$/мес. |
| 2. Возможности парсинга  | Универсальный  | 1. Парсинг ссылок.<br>2. Парсинг с помощью WBApP.<br>3. Парсинг статей с картинками и файлами.<br>4. Парсинг XML-карт сайтов. | Универсальный   |   |
| 3. Возможности выгрузки  | CSV/TXT/База данных/Excel;<br>WordPress;<br>DLE;<br>Webasyst;<br>Joomla;<br>и др.                          | CSV (с любыми заданными столбцами), htm, txt, php, MySQL  | Excel, JSON, HTML, базы данных  | API<br>CSV/ Excel<br>Google-таблицы<br>Tableau  |
| 4. Основная применимость | Универсальный (есть наборы настроек для конкретных целей – интернет-магазины, соцсети, порталы объявлений. | Универсальный - подходит для парсинга блогов, соц сетей, страниц с новостями  | Универсальный (в т.ч. сканирует все типы социальных сетей. Быстрые и точные результаты)<br>Массовое извлечение полей данных.  | Универсальный, поддерживается большинством систем.  |

### Литература

1. Коростелев А.В., Мартышкин А.И. Возможность автоматизации процесса поиска целевой аудитории «Вконтакте» с использованием данных об оценках пользователей другой социальной сети // XXI век: итоги прошлого и проблемы настоящего плюс. 2020. № 2 (50). С. 149-157.
2. Алейников Д.П., Руденко М.Б. Анализ цифрового контента социальных сетей на предмет выявления оперативно значимой информации // Стратегическое развитие

- системы МВД России. Сборник трудов Международной научно-практической конференции: состояние, тенденции, перспективы. 2020. С. 40-44.
3. Калашник М.В., Ковалёв А.В., Сиганов И.Д. Анализ городского пространства по активности горожан в социальных сетях // Молодёжь третьего тысячелетия». Сборник трудов региональной Конференции. Омск. 2019. С. 252-255.
  4. Беген П.Н., Низомутдинов Б.А., Тропников А.С. Классификатор объектов городского хозяйства для данных из социальных сетей // Научный сервис в сети Интернет: труды XXII Всероссийской научной конференции (21-25 сентября 2020 г., онлайн). М.: ИПМ им. Келдыша. 2020. С. 101-115.
  5. Просветов В.Н., Конева Н.Е. Анализ методов и средств автоматизации процессов обработки данных веб-сайтов // Евразийское Научное объединение. 2019. №1-2 (47). С. 89-94.



**Рюмина Елена Витальевна**

Год рождения: 1991  
Санкт-Петербургский институт информатики и автоматизации  
Российской академии наук,  
программист,  
Университет ИТМО,  
факультет информационных технологий  
и программирования,  
студент группа №М41211,  
направление подготовки: 09.04.02 – Информационные  
системы и технологии,  
e-mail: riumina\_ev@mail.ru



**Дресвянский Денис Владиславович**

Год рождения: 1995  
Университет ИТМО,  
факультет информационных технологий и программирования,  
спирант группа №7717,  
направление подготовки: 05.13.11 – Математическое  
и программное обеспечение вычислительных машин,  
комплексов и компьютерных сетей,  
e-mail: denis.dresvyanskiy@uni-ulm.de



**Карпов Алексей Анатольевич**

Год рождения: 1978  
Санкт-Петербургский институт информатики и автоматизации  
Российской академии наук,  
д.т.н., доцент, г.н.с.,  
e-mail: karpov@iias.spb.su

УДК 004.932.2

**ВЛИЯНИЕ МЕТОДОВ СНИЖЕНИЯ  
СКОРОСТИ НА ЭФФЕКТИВНОСТЬ СИСТЕМ  
РАСПОЗНАВАНИЯ ЭМОЦИЙ ПО МИМИКЕ ЛИЦА**

**Е.В. Рюмина, Д.В. Дресвянский**

**Научный руководитель – д.т.н., доцент, г.н.с. Карпов А.А.**

Работа выполнена при финансовой поддержке Российского научного фонда РФФ №18-11-00145 «Разработка и исследование интеллектуальной системы для комплексного паралингвистического анализа речи».

**Аннотация**

В работе рассмотрены различные методы снижения скорости обучения для задачи распознавания эмоций по мимике лица человека с использованием сверточной нейронной сети ResNet-50. Осуществлен анализ влияния различных оптимизаторов при разных начальных скоростях обучения. Исследование проводилось на крупномасштабном наборе данных изображений выражений лиц AffectNet. Результаты экспериментальных исследований показывают, что при правильном подборе начальной

скорости обучения, оптимизатора и метода снижения скорости обучения можно повысить точность системы распознавания эмоций по мимике лица человека на 7,43%.

### **Ключевые слова**

Распознавание эмоций, мимика лица, ResNet-50, Time-based, Piecewise Constant, Cosine Annealing.

### **Введение**

Эмоция – это психофизиологический процесс, который отражает внутренние намерения человека. Возможность создания искусственного эмоционального интеллекта, способного адаптироваться под эмоциональные состояния пользователей, привлекает не только научных исследователей, но и пользуется большим спросом в таких областях, как медицина, психология, безопасность и др. Таким образом, эмоции играют важную роль в жизнедеятельности человека, а системы распознавания эмоций имеют как научную, так и социальную значимость, которую невозможно переоценить.

Эмоции могут выражаться как при вербальном (например, речь и письмо), так и при невербальном (например, мимика лица и жесты) каналах общения. Несомненно, мимика лица является одним из основных источников информации в межличностном общении. Однако существующие системы распознавания эмоций по мимике лица человека все еще не обладают возможностями, необходимыми для построения надежной системы, способной работать в неконтролируемых условиях [1]. Современные подходы к распознаванию эмоций по мимике лица человека строятся на основе сверточных нейронных сетей [1, 2]. Разница в подходах, основанных на сверточных нейронных сетях, заключается в их архитектурах: в зависимости от архитектуры применяемой сети достигается различная точность распознавания. Однако не только подбор архитектуры сети способен улучшить точность распознавания, но и подбор оптимизатора, скорости обучения, метода снижения скорости обучения и настройка других параметров также способствуют повышению эффективности распознавания. Таким образом, в данной работе осуществлен аналитический обзор влияния различных методов снижения скорости обучения при различных оптимизаторах и начальных скоростях обучения на точность распознавания эмоций по мимике лица человека.

### **Методы снижения скорости обучения**

Скорость обучения является одним из наиболее важных гиперпараметров при машинном обучении с помощью глубоких нейронных сетей. Скорость обучения контролирует сходимость сети к точке глобального минимума, но при неправильном ее подборе сеть может «застрять» в точке локального минимума, и, следовательно, система не достигнет максимальной эффективности решения поставленной задачи. Для достижения глобального минимума необходимо применять методы снижения/увеличения скорости в процессе обучения сети. В научной литературе существуют примеры исследовательских работ, продемонстрировавших эффективность применения таких методов [4].

В рамках данного исследования были рассмотрены 3 метода снижения скорости обучения, а именно: Time-based, Piecewise Constant, Cosine Annealing и традиционный метод обучения сети при фиксированной скорости обучения на протяжении всего процесса обучения (рис. 1а). Time-based – это метод снижения скорости обучения, при котором она уменьшается в зависимости от текущей эпохи и вычисляется по формуле (1). На рис. 1б представлен график скорости обучения, где начальная скорость была установлена на 0,001 с постепенным понижением скорости, равным 0,001/30, где 30 – количество эпох.

$$lr = lr_{start} * \frac{1}{1 + \frac{lr_{start} * epoch_{curr} * N}{epoch * batch\_size}}, \quad (1)$$

где  $lr_{start}$  – начальная скорость обучения;  $epoch$  – общее количество эпох обучения;  $epoch_{curr}$  – номер текущей эпохи в момент обновления скорости обучения;  $N$  – общее количество экземпляров в обучающей выборке;  $batch\_size$  – размер батча.

Cosine Annealing с холодным перезапуском – это метод скорости обучения, заключающийся в начале с большой скорости обучения, относительно быстрым ее снижением до установленного минимального значения и, затем, возвратом к первоначальному значению. Обновление скорости обучения происходит по формуле (2). На рис. 1с, в качестве примера, представлен график изменения скорости обучения по методу Cosine Annealing с начальной скоростью обучения, равной 0,001 и минимальным возможным значением 0,0001. Количество циклов перезапуска – 5 в течение 30 эпох.

$$lr = \frac{lr_{start}}{2} * \left( \cos \left( \frac{\pi * \text{mod}(epoch_{curr} - 1, \lfloor \frac{epoch}{M} \rfloor)}{\lfloor \frac{epoch}{M} \rfloor} \right) + 1 \right), \quad (2)$$

где  $\cos()$  – косинус числа;  $\text{mod}()$  – остаток от деления;  $M$  – количество циклов.

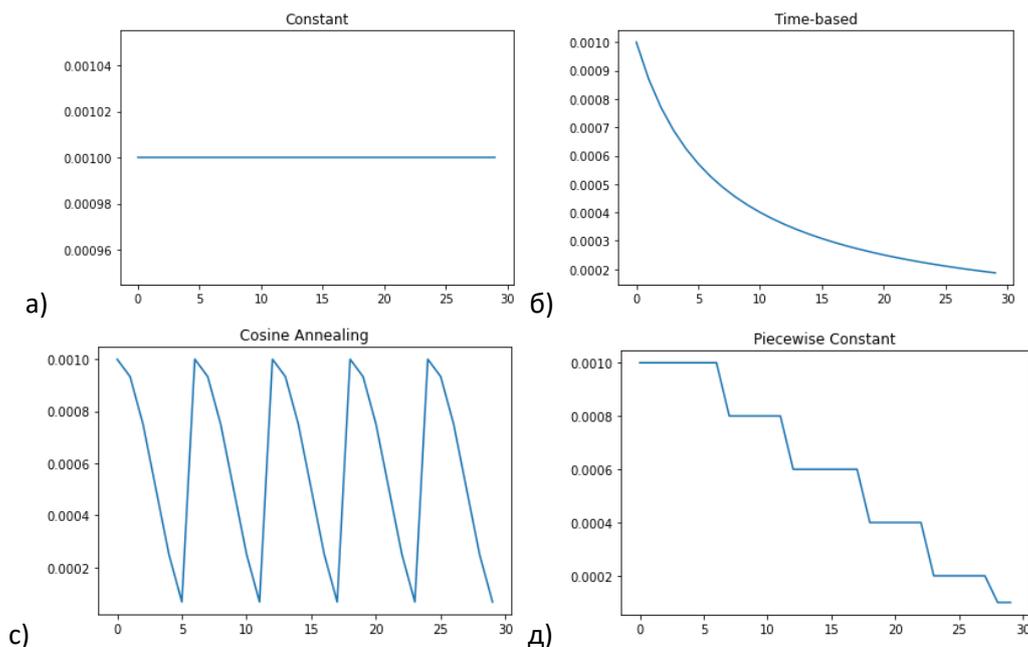


Рис. 1. Графики исследуемых скоростей обучения: а) скорость фиксирована на всем этапе обучения б) Time-based, в) Cosine Annealing д) Piecewise Constant

Piecewise Constant – это метод снижения скорости обучения, при котором скорость изменяется при заданных значениях итераций (мини-батчей) на заданный параметр скорости. На рис. 1д представлен график изменения скорости обучения, где начальная скорость была установлена на 0,001, затем, на 6 эпохе скорость обучения была уменьшена до 0,0008, на 11 эпохе - до 0,0006, и так далее до последнего снижения на 27 эпохе - до 0,0001.

Таким образом, в работе исследуются графики обучения (см. рис. 1), а также рассматриваются различные начальные скорости обучения 0,01, 0,001, 0,0001 и оптимизаторы Adam и SGD.

### Исследовательские данные

Для исследования был выбран крупномасштабный набор данных изображений выражений лиц AffectNet [1]. Изображения собраны в сети Интернет и охватывают лица разных национальностей, возрастов, пола, степени окклюзий, освещенности и фона. Имеется большая как внутриклассовая, так и межклассовая вариативность в данных, что делает данный набор наиболее подходящим для текущего исследования. На рис. 2 представлены примеры из набора данных AffectNet.



Рис. 2. Примеры изображений из набора данных AffectNet

Создатели AffectNet предлагают изображения с разметкой на 8 категориальных эмоций (злость, счастье, грусть, отвращение, удивление, презрение, страх и нейтральное состояние), однако эмоция презрение не является базовой по мнению Пола Экмана [3], поэтому в рамках данного исследования не рассматривается. В табл. 1 представлено распределение экземпляров по категориальным эмоциям в обучающем наборе. Предлагается также проверочный набор, содержащий по 500 экземпляров каждого класса. Стоит отметить, что, так как количество экземпляров для каждого класса в тестовой выборке одинаковое, то нет необходимости в использовании такой метрики, как невзвешенная средняя полнота (Unweighted Average Recall), поэтому в рамках данного исследования все результаты измеряются с помощью метрики точность (Accuracy).

Таблица 1

#### Распределение экземпляров по эмоциональным категориям

| Эмоция             | Нейтраль. | Счастье | Печаль | Удивление | Страх | Отвращение | Злость |
|--------------------|-----------|---------|--------|-----------|-------|------------|--------|
| Кол-во экземпляров | 74 874    | 134 415 | 25 459 | 14 090    | 6 378 | 3 803      | 24 882 |

Как можно заметить из табл. 1, в обучающем наборе данных имеется сильный дисбаланс классов. Данную проблему можно решить с помощью, например, аугментации данных или взвешивания классов. Далее в разделах Предварительная обработка данных и Предлагаемая нейронная сеть, будут подробно описаны методы, использовавшиеся в рамках данного исследования для решения проблемы дисбаланса классов.

#### Предварительная обработка данных

Набор данных AffectNet содержит обрезанные изображения выражений лиц, поэтому изображения в дополнительном детектировании лица не нуждаются. Однако, так как изображения представлены в формате RGB с диапазоном пикселей от 0 до 255, требуется нормализация данных. Так, в исследовании используется предобученная на наборе данных VGGFace2 [5, 6] сверточная нейронная сеть ResNet-50, на вход которой

необходимо подавать центрированные значения пикселей изображения для каждого RGB канала согласно следующему: 1) из каждого пикселя канала Red отнимается значение 91,4953; 2) канал Green – 103,8827; канал Blue – 131,0912.

Как было сказано ранее, в наборе данных AffectNet наблюдается дисбаланс классов, поэтому в исследовании применяется две техники аугментации данных:

1. Добавление экземпляров из набора данных FER2013 [6] для миноритарных классов (для всех классов, кроме счастья и нейтральное состояние). На рис. 3 представлены примеры изображений из набора данных FER2013.



Рис. 3. Примеры изображений из набора данных FER2013

Такая техника позволяет увеличить количество экземпляров в обучающем наборе для следующих эмоций: страх до 10475 экземпляров (на 39,11%); печаль – 30289 (15,95%), злость – 28877 (13,83%), удивление – 17261 (18,37%), отвращение – 4239 (12,17%).

2. Случайные аффинные преобразования (сжатие, растяжение, отображение по горизонтали, вращение, искажение) и изменение контрастности изображения. На рис. 4 представлены примеры применения случайных аффинных преобразований и изменения контрастности относительно оригинального изображения. Данная техника вносит изменения в обучающие данные в процессе самого обучения на каждой эпохе, при этом не увеличивая количество экземпляров в обучающем наборе.



Рис. 4. Примеры применения случайных аффинных преобразований и контрастности

Применение описанных выше техник аугментации данных позволяет увеличить робастность и точность системы распознавания эмоций по мимике лица человека.

### Предлагаемая нейронная сеть

В качестве алгоритма машинной классификации и извлечения векторов признаков использовалась предобученная сверточная нейронная сеть ResNet-50 [7]. Применение предобученных нейронных сетей на одной задаче для решения новой задачи называется трансферным обучением. Такой подход позволяет сети гораздо быстрее обучиться и требует меньше обучающих данных.

Так, веса нейронной сети ResNet-50 были использованы в качестве отправной точки для дообучения сети на решение задачи распознавания эмоций по мимике лица человека. Однако, так как нейронная сеть была предобучена на решение задачи распознавания лиц (8 631 классов), то последний классифицирующий слой

замораживался, затем добавлялись новые слои: 1) слой глобального усредненного пулинга; 2) слой с случайным гауссовским шумом при стандартном отклонении 0,1 (применяется только во время обучения сети); 3) полносвязный слой с количеством выходных нейронов 512 и 12 регуляризацией 0,0001; 4) слой со случайным прореживанием нейронов с 50% вероятностью; 5) полносвязный слой с количеством выходных нейронов, равным 7 и с функцией активации Softmax.

Также для обращения внимания нейронной сети на миноритарные классы использовалось обратно пропорциональное частотам взвешивание классов, которое рассчитывается по формуле (3):

$$\text{вес}_i = \frac{\text{Общее кол-во примеров}}{\text{Кол-во примеров для класса } i * \text{Кол-во классов}}, \quad (3)$$

где  $i$  – индекс класса от 0 до 6.

### Экспериментальные исследования

В табл. 2 представлены значения точности, полученные в ходе использования описанной ранее нейронной сети при различных методах снижения скорости обучения, начальных скоростях обучения и оптимизаторах.

Как можно заметить из табл. 2, минимальное значение точности (57,14%) было достигнуто при методе снижения скорости обучения Time-based, оптимизаторе Adam и начальной скорости 0,001. Максимальное же значение точности (64,57%) было достигнуто при методе снижения скорости обучения Cosine Annealing, оптимизаторе SGD и начальной скорости 0,0001. Кроме того, с помощью метода снижения скорости обучения Cosine Annealing удалось достигнуть значения точности выше 64% при использовании обоих оптимизаторах, тогда как при других методах снижения скорости обучения этого не наблюдается. Также можно заметить, что, в целом, выбор оптимизатора оказывает меньшее влияние на конечную точность классификации по сравнению с выбором начальной скорости обучения. Кроме того, примерно равные значения точности для оптимизаторов Adam и SGD достигаются в том случае, если начальная скорость обучения при SGD в 10 раз выше по сравнению с начальной скоростью обучения для оптимизатора Adam.

Таблица 2

### Экспериментальные результаты

| Метод снижения скорости обучения | Оптимизатор | Начальная скорость обучения |        |         |
|----------------------------------|-------------|-----------------------------|--------|---------|
|                                  |             | 0,001                       | 0,0001 | 0,00001 |
|                                  |             | Точность, %                 |        |         |
| Constant                         | Adam        | 57,4                        | 62,74  | 63,86   |
|                                  | SGD         | 62,4                        | 64,26  | 62,86   |
| Time-based                       | Adam        | 57,14                       | 61,67  | 64,09   |
|                                  | SGD         | 62,94                       | 63,94  | 62,74   |
| Piecewise Constant               | Adam        | 59,17                       | 62,17  | 63,89   |
|                                  | SGD         | 62,94                       | 63,37  | 62,91   |
| Cosine Annealing                 | Adam        | 59,54                       | 62,57  | 64,17   |
|                                  | SGD         | 62,68                       | 64,57  | 62,97   |

### Выводы

В работе были рассмотрены 3 метода снижения скорости обучения, а именно Time-based, Piecewise Constant, Cosine Annealing и традиционный метод обучения сети при фиксированной скорости на протяжении всего процесса обучения. При проведении экспериментальных исследований каждый из методов был обучен с использованием различных оптимизаторов и начальных скоростей обучения. Обучение и тестирование проводилось на крупномасштабном наборе данных изображений выражений лиц AffectNet с использованием сверточной нейронной сети ResNet-50. С помощью предлагаемого подхода для 7 категориальных эмоций (презрение не рассматривалось) была достигнута точность распознавания 64,57%. Кроме того, полученные результаты позволяют сделать вывод, что применение различных методов снижения скоростей, а также подбор значения начальной скорости обучения и оптимизатора позволяют получить абсолютный прирост в значении точности для задачи распознавания эмоций по мимике лица человека на 7,43%.

### Литература

1. Mollahosseini A, Hasani B, Mahoor MH. Affectnet: A database for facial expression, valence, and arousal computing in the wild // IEEE Transactions on Affective Computing. 2017. Vol. 10. № 1. P.18-31, 5.
2. Szegedy C., Ioffe S., Vanhoucke V., Alemi. A. Inception-v4, Inception-Resnet and the impact of residual connections on learning // arXiv preprint arXiv:1602.07261. 2016.
3. Ekman P. Are there basic emotions? // Psychol Rev. 1992. Vol. 99. № 3. P. 550-553. doi: 10.1037/0033-295x.99.3.550. PMID: 1344638.
4. Mishra P., Sarawadekar K. Polynomial learning rate policy with warm restart for deep neural network // IEEE Region 10 Conference (TENCON). – 2019. – P. 2087-2092.
5. Cao Q, Shen L, Xie W, Parkhi OM, Zisserman A. Vggface2: A dataset for recognising faces across pose and age // Proceedings of the 13th IEEE International Conference on Automatic Face & Gesture Recognition (FG 2018). – 2018. – P. 67-74.
6. Challenges in Representation Learning Facial Expression Recognition Challenge [Электронный ресурс]. <http://www.kaggle.com/c/challenges-in-representation-learning-facialexpression-recognition-challenge>, свободный. Яз. англ. (дата обращения 07.02.2021).
7. He K., Zhang X., Ren S., Sun J. Deep residual learning for image recognition // Proceedings of the IEEE conference on computer vision and pattern recognition. 2016. P. 770-778.



**Скрипачев Иван Сергеевич**  
Университет ИТМО,  
Институт дизайна и урбанистики,  
студент группы № С42041,  
направление подготовки: 09.04.03 – Прикладная информатика,  
e-mail: skripacheviv@gmail.com



**Митягин Сергей Александрович**  
Университет ИТМО,  
Институт дизайна и урбанистики,  
к.т.н.,  
e-mail: mityagin@itmo.ru

УДК 004.03

**РАЗРАБОТКА МЕТОДА ПЛАНИРОВАНИЯ  
ОПТИМАЛЬНОГО РАЗМЕЩЕНИЯ  
ОБЪЕКТОВ НАРУЖНОГО ОСВЕЩЕНИЯ  
И.С. Скрипачев  
Научный руководитель – к.т.н. С.А. Митягин**

Работа выполнена в рамках темы НИР «Разработка метода планирования оптимального размещения объектов наружного освещения».

**Аннотация**

Основной проблемой при планировании наружного освещения является оптимальное размещение осветительных объектов. В данной работе были проанализированы существующие методы размещения объектов в пространстве. На основе анализа был разработан метод размещения объектов освещения и экспериментально исследован на тестовых областях.

**Ключевые слова**

Наружное освещение, планирование, оптимизация, комбинаторика, машинное обучение, искусственный интеллект, генетический алгоритм, сверточная нейронная сеть, обучение с подкреплением, Unity3d.

В общем виде задачу о размещении объектов освещения можно описать следующим образом. Требуется разместить  $N$  источников освещения типа  $T$  с параметрами  $T_p$  на территории  $M$  размером  $X$  на  $Y$ . На территории имеется определенный набор объектов типа  $U$ . Целью размещения является оптимизация совокупности параметров  $(k_1, k_2, \dots, k_i)$  для двух вариантов решения задачи:

1. Максимизация площади освещённости исследуемой территории.

$$f(M, g(N), g(T_p)) \overrightarrow{max} \sum_{i=0}^{num(U)} S(U_i),$$

где  $g(x)$  – функция ограничения параметра,  $S$  – площадь освещенности,  $min(x)$  – количество  $x$ .

2. Минимизация количества источников освещения (экономичность).

$$f(M, g(N), g(T_p)) \overline{min} N,$$

где  $g(x)$  – функция ограничения параметра  $x$ ,



Рис. 1. Графическое представление поставленной задачи

Параметр количества источников  $N$  является оптимизируемым и требует минимизации с целью увеличения экономичности либо максимизации с целью увеличения площади освещенности (рис. 1). Данный параметр может иметь ограничения минимального и максимального количества, установленные входными данными задачи.

Параметр типа источника освещения  $T$  имеет ряд параметров  $T_p$ :

- форма пятна освещения;
- размер пятна освещения.

Данный параметр также является оптимизируемым по отношению к экономичности и площади освещенности. Имеет ряд ограничений, связанных со строительными правилами и нормами.

Параметры  $X$  и  $Y$  задают размеры исследуемой территории. В задаче они должны варьироваться в определенном допустимом интервале (ограничивается выбранным алгоритмом). Данные параметры являются входными и не оптимизируются.

Параметр  $U$  описывает тип объекта, который следует осветить (например, дорожное покрытие, пешеходная тропа, площадь и т. д.). Данный параметр вводит ряд ограничений на выбор источника освещения, его характеристик, а также на тип устанавливаемой опоры. Каждый объект имеет свою форму  $S$  и размер  $R$ . Форма является замкнутой фигурой в двумерном пространстве.

Совокупностью параметров  $(k_1, k_2, \dots, k_i)$  выступают как вышеописанные параметры, так и параметры, вычисляемые в ходе решения задачи. В качестве таких параметров могут быть:

- координаты источника  $L_i$  в пространстве  $(x_{li}, y_{li})$ ;
- тип и характеристики опоры.

Существует ряд традиционных (классических) подходов и методов решения задач размещения объектов. Среди которых «Задача о размещении объектов», «Задача о рюкзаке», «Задача упаковки» (методы комбинаторики) [1]. Однако все эти методы имеют ряд недостатков, а именно [2]:

- сложность представления входных данных,
- невозможность определения местоположения объекта,

– геометрические ограничения заполняемого пространства.

В качестве решения поставленной задачи был исследован подход с искусственным интеллектом [3]. Для этого были объединены два инструмента: сверточная нейронная сеть и генетический алгоритм [4]. В качестве входных данных были взяты примитивные карты, представленные матрицами со значением яркости пикселя, с исследуемой территорией. Каждая карта имеет размер 10 на 10. Всего на карте 20 белых пикселей, требующих освещения. Также карты были нормализованы, поэтому черный пиксель имеет значение 0, а белый – 1. Примеры этих карт представлены на рис. 2.

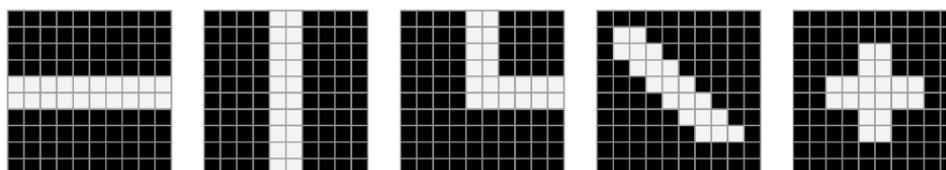


Рис. 2. Пример карт исследуемых территорий

Для тестирования использовались различные варианты архитектуры нейронной сети. Пример одной из оптимальных архитектур сверточной нейронной сети, полученной в ходе экспериментов, представлен на рис. 3.

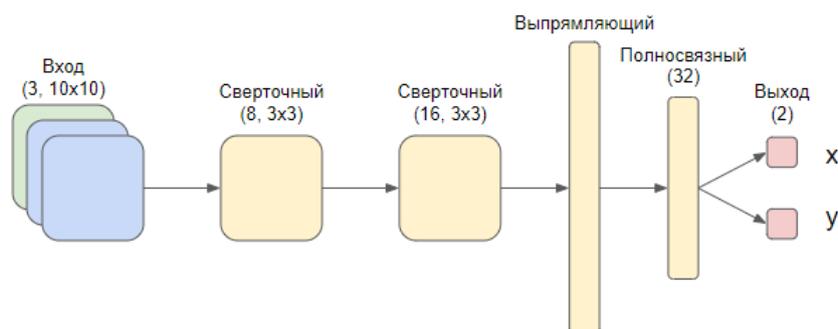


Рис. 3. Схема сверточной нейронной сети

Для оценки качества действий нейронной сети была выбрана следующая фитнес-функция.

$$f(x, y) = \begin{cases} +1, & M(x, y) = 1 \\ -1, & M(x, y) = 0 \end{cases},$$

где  $M(x, y)$  – значение яркости пикселя с координатами  $x$  и  $y$  на карте.

При успешном освещении нужного пикселя система поощрялась на +1 балл. При неудачном выборе пикселя система штрафовалась на -1 балл. Число итераций было ограничено максимальным числом пикселей, требующих освещения – то есть 20. Максимальная возможная оценка равняется +20 баллам. Минимальная -20 баллам. Данная оценка нужна для выбора лучших вариантов среди нейронных сетей с помощью генетического алгоритма. Максимальная достигнутая оценка составила -10 баллов. Это также можно считать неудовлетворительным результатом. Точность такой сети составила 25%. На рис. 4 представлен график процесса обучения.

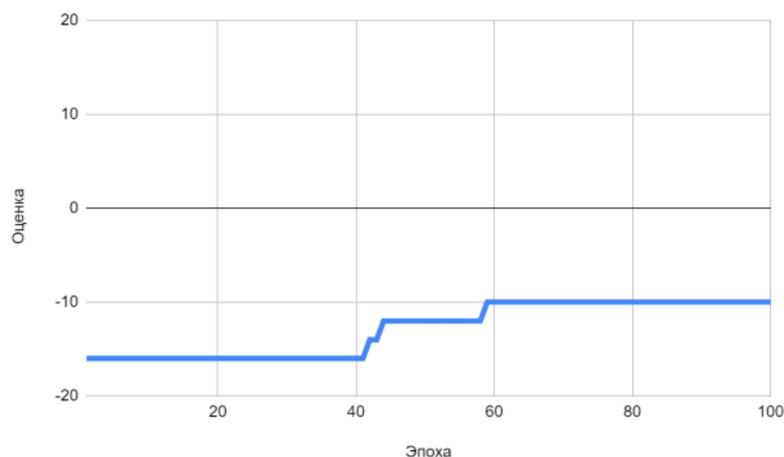


Рис. 4. График процесса обучения

Результаты исследования показали, что данный подход является неприменимым для поставленной задачи размещения объектов освещения даже в упрощенном виде.

Второй метод был разработан на основе комбинаторной оптимизации и физических свойствах объектов. Его суть заключается в заполнении некоторой ограниченной геометрической формы шарами. Путем перебора различных комбинаций генераторов, линейных сдвигов и параметров шаров достигается одно из условий, описанных в поставленной задаче. Сам процесс позаимствован из реального мира и напоминает укладку шаров в коробку либо укладку бильярдных шаров в треугольник перед началом партии.

Шар представляет собой условное пятно распространения освещения, определенного радиуса. На каждый шар действуют физические силы (например, гравитация), симулируемые с помощью игрового движка Unity3D [5]. Также симулируется система столкновений, реализуемая за счет коллайдеров, которыми обладает шар.

Коробка – это исследуемая область определенного типа на входной территории. В качестве такой области может выступать участок дороги, тропинки, точки отдыха и т. д. Подразумевается, что входная карта имеет координаты вершин объектов на ней. Затем алгоритм реконструирует данную область в 3D пространстве на движке Unity3d (рис. 5).

Генератор – это условный объект для генерации шаров. Для симуляции используется два вида генераторов: «Для генерации одиночных шаров» и «Шаблонный генератор шаров».

Алгоритм разработанного метода представлен ниже.

1. Генерация исследуемой области для каждого объекта, который требуется осветить.

2. Для каждого генератора (шаблона):

а. Для каждого радиуса шара (изменяется с заданным шагом):

б. Для каждого смещения генератора (изменяется с заданным шагом):

– Сбрасываем шары в коробку.

– Ждем, когда они самоорганизуются.

– Удаляем шары, которые не коснулись нижней поверхности коробки.

– Повторяем N раз цикл начиная со сброса шаров до текущего шага.

– Получаем оценку решения (площадь покрытия и количество источников) и сохраняем координаты.

3. Выбираем лучшее решение по критерию площади покрытия и по критерию количества источников.

В процессе тестирования данного подхода было создано несколько эвристик, значительно улучшающих решение (рис. 5).

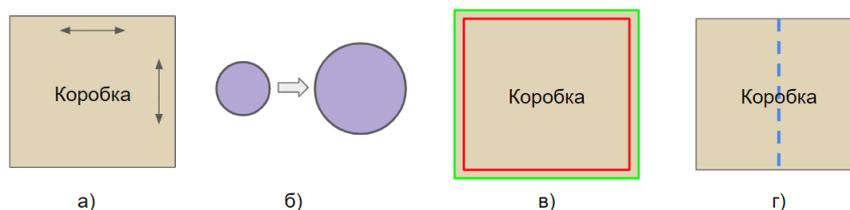


Рис. 5. а – Эффект «Встряски»; б – Плавное увеличение шаров;  
в – Расширение границ коробки; г – Вариативное разбиение коробки на подобласти

Данный метод был протестирован на нескольких тестовых областях, имитирующих реальные пространства. В одном из экспериментов тестировалась Г-образная область. На рис. 6 представлены результаты работы метода. Полученные результаты удовлетворяют условиям задачи. Оба варианта имеют практически одинаковый процент покрытия освещением  $\sim 82\%$ . Однако в первом случае требуется 20 осветительных приборов с небольшим радиусом освещенности, а во втором 5 мощных приборов. Шары расположены равномерно и структурировано. Для данного эксперимента были применены все эвристики кроме деления области на подобласти.

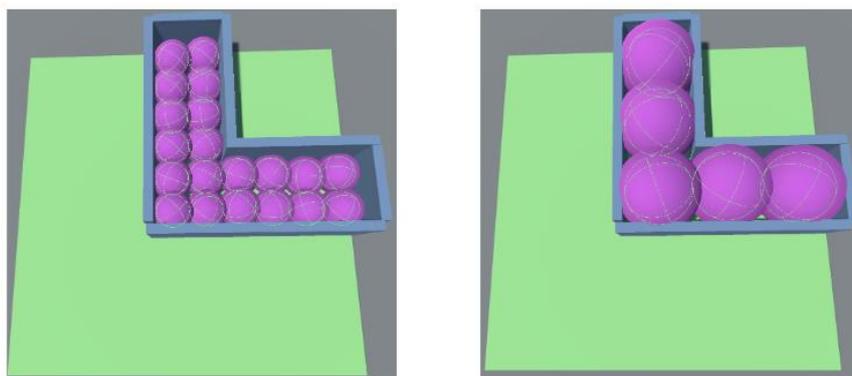


Рис. 6. Слева – максимизация покрытия. Справа – минимизации количества источников

Также интересные результаты были получены для О-образной формы. Визуально их можно оценить на рис. 7. Полученные результаты удовлетворяют условиям задачи. В данном эксперименте показатели покрытия оказались разными. Для варианта с максимизацией покрытия -  $\sim 80\%$ , а для варианта с минимизацией количества источников -  $\sim 75\%$ . Количество источников освещения в первом случае – 19, во втором – 4. Как и в предыдущем эксперименте, видно, что объекты расположены равномерно и структурировано.

В данном исследовании были проанализированы существующие классические подходы и методы размещения объектов. Анализ позволил выявить недостатки применимости данных подходов к поставленной задаче.

Поэтому были разработаны два решения для размещения объектов освещения. Первый метод основан на машинном обучении с подкреплением и показал неудовлетворительные результаты. Точность полученных решений не превышала 25% для упрощенных моделей.

Второй метод был разработан на основе комбинаторной оптимизации и физических свойствах объектов. Случайные процессы при самоорганизации объектов освещения внутри исследуемой области являются как его преимуществом, так и

недостатком. С одной стороны, получаются уникальные решения, которые невозможно получить при классических подходах, с другой стороны, нет гарантий получения точного решения. Результаты исследования показали, что данный подход применим для поставленной задачи размещения объектов освещения. Ряд разработанных эвристик помог достичь максимальной площади покрытия освещением в 80%.

### Литература

1. Schrijver A. Algorithms and Combinatorics. – Springer-Verlag Berlin Heidelberg, 2003. 1881 с.
2. Левитин А.В. Алгоритмы. Введение в разработку и анализ. М.: Вильямс. 2006. 576 с.
3. Рашка С. Python и машинное обучение. М.: ДМК Пресс. 2017. 418 с.
4. Wikipedia. Genetic algorithm [Электронный ресурс] / Режим доступа: [https://en.wikipedia.org/wiki/Genetic\\_algorithm/](https://en.wikipedia.org/wiki/Genetic_algorithm/), свободный. Загл. с экрана. Яз. англ.
5. Unity [Электронный ресурс] / Режим доступа: <https://unity.com/ru/>, свободный. Загл. с экрана. Яз. англ.



**Соловьева Валентина Викторовна**

Год рождения: 1984

Университет ИТМО,

Институт дизайна и урбанистики,

студент группы №7738,

направление подготовки: 05.13.10 – Управление

в социальных и экономических системах,

e-mail: valentina.soloveva@itmo.ru

УДК 338.27

## **АНАЛИЗ ФАКТОРОВ, ВЛИЯЮЩИХ НА СТОИМОСТЬ ЭКОНОМИИ ВРЕМЕНИ В ПУТИ ЖИТЕЛЕЙ ГОРОДА**

**В.В. Соловьева**

**Научный руководитель – к.т.н. С.А. Митягин**

### **Аннотация**

В работе рассмотрены стандартные применяемые методы определения стоимости времени в пути, определена проблематика неточности данных методов. Приведены конкретные кейсы, используемые автором для изучения стоимости времени в пути в российских городах. Предложены новые факторы, влияющие на стоимость экономии времени в пути жителей городов.

### **Ключевые слова**

Стоимость времени в пути, стоимость экономии времени в пути, готовность платить за сэкономленное время, платная транспортная инфраструктура, платные дороги.

Стоимость времени в пути (Value of Travel Time, VTT) – это стоимость времени, затраченного на поездку. Стоимость экономии времени в пути (Value of Travel Time Savings, VTTS) – стоимость, которую человек готов заплатить за то, чтобы доехать быстрее, то есть за сэкономленное время в поездке.

Стоимость экономии времени в пути – это величина, которая позволяет перевести время, потраченное на поездку человека, в денежные единицы. Аналогом этой величины можно назвать готовность платить за сэкономленное время в пути. Готовность платить за минуту сэкономленного времени в пути является критически важным показателем в транспортном планировании, когда речь идёт о введении платной транспортной инфраструктуры (платная дорога, платная парковка или плата за общественный транспорт) и расчёте тарифов на неё.

Эта величина не может быть одинаковой для всех людей. Она будет различной в зависимости от места жительства, транспортного поведения и уровня дохода. Исследования показывают, что для оценки стоимости времени в пути важны цель поездки, вид транспорта, доход и преодолеваемое расстояние [1, 2]. Данные многих исследований свидетельствуют о том, что доход оказывает существенное влияние на стоимость времени в пути, но возраст, семейное положение и род занятий также оказывают влияние, например, семьи выбирают быструю альтернативу, даже если она дороже [3]. Для различных целей исследования может быть необходимо определить среднюю стоимость времени в пути для всех жителей или посчитать отдельный показатель для разных групп (по уровню дохода, по цели поездки и др.)

Упрощённо за стоимость времени в пути принимается показатель средней заработной платы, приведённый к минуте. Однако такой подход не учитывает большое

количество факторов. Так, время, проведённое на работе, время, посвящённое досугу, и время в пути имеют для человека разную ценность [4, 5].

Один из наиболее популярных методов определения стоимости времени в пути – социологические исследования. Респонденту предлагают набор гипотетических ситуаций, в которых он может сэкономить определённое количество времени в пути за определённую сумму. При делении этой суммы на время получают стоимость времени в пути в руб./минуту.

Так была рассчитана стоимость времени автомобилистов из Владивостока, где был проведён социологический опрос среди водителей<sup>1</sup>. Для выявления стоимости экономии времени в пути респондентам был задан ряд вопросов с гипотетическим выбором, касающийся его последней поездки на автомобиле: «Представьте, что для этой поездки вы могли бы использовать платную дорогу, которая уменьшила бы время в пути на 10 минут. Какова максимальная сумма, которую вы готовы заплатить за такую экономию?». Средняя стоимость времени в пути составила 8,2 руб./минуту.

Минус этого метода – выявление субъективной оценки готовности платить (так называемые заявленные предпочтения), а субъективная оценка при гипотетическом выборе может отличаться от реального выбора, который сделает человек в заданных условиях.

Объективно показатель стоимости времени в пути можно получить, используя объективные данные о перемещениях жителей. Например, в случае имеющейся платной транспортной инфраструктуры и больших данных о перемещении жителей можно выявить, сколько заплатил человек за сэкономленное время и в каком случае принял то или иное решение.

Таким образом, было проведено исследование стоимости времени в пути в Санкт-Петербурге<sup>2</sup> – средний показатель составил 13,3 руб./минуту. Однако этот метод зачастую неприменим из-за отсутствия таких объективных данных.

Все описанные выше методы рассматривают стоимость экономии времени в пути как единый показатель – готовность заплатить определённую сумму за минуту сэкономленного времени. То есть предполагается, что при совершении поездки водитель каждый раз принимает решение на основании потенциальных временных и финансовых затрат. При этом принимается, что каждая минута сэкономленного времени стоит одинаково.

Однако на показатель стоимости времени в пути могут также влиять и другие факторы, а общая стоимость сэкономленного времени может распределяться нелинейно.

Можно предположить, что общее время поездки может влиять на стоимость времени в пути. Например, экономия в 10 минут более существенна для коротких поездок, где она составляет значительную часть общего времени в пути, чем для длинных поездок, где доля такой экономии ниже. Поэтому в упомянутом выше опросе жителей Владивостока тем респондентам, чья поездка была длиннее 30 минут, дополнительно был задан вопрос про экономию 20 минут, а тем респондентам, чья поездка длилась более 60 минут – ещё один вопрос про экономию 30 минут.

В результате стоимость, которую респондент был готов заплатить, не увеличивалась линейно: она либо оставалась такой же, как при ответе на первый вопрос об экономии 10 минут, либо увеличивалась, но непропорционально предложенному количеству сэкономленных минут. Таким образом, можно сделать

---

<sup>1</sup> Опрос был проведён автором и коллегами в 2020 году среди автомобилистов Владивостока, выборка составила  $\geq 1000$  респондентов.

<sup>2</sup> Исследование проводилось автором и коллегами в 2019 году с использованием социологического опроса автомобилистов Санкт-Петербурга (выборка  $\geq 1000$  респондентов) и прогнозных данных навигаторов о времени в пути по платной и бесплатным дорогам.

вывод, что с увеличением длительности поездки стоимость экономии времени в пути за одну минуту падает.

Это можно объяснить тем, что человек может быть готов заплатить за экономию времени как таковую, а количество сэкономленных минут играет при этом меньшую роль. К тому же, вероятно, в сумму, которую человек готов заплатить за экономию времени в пути, кроме стоимости времени, входят также иные параметры, такие как надёжность времени в пути, комфорт или другие.

Надёжность времени в пути определяет вероятность отклонения от запланированной длительности пути: если надёжность высокая, человек затратит ровно то время на поездку, которое он ожидает.

Для оценки стоимости надёжности времени в пути респондентам того же опроса во Владивостоке предлагалось сделать гипотетический выбор между платной поездкой, длящейся  $32 \pm 1$  минуты, и бесплатной поездкой длительностью  $46 \pm 12$  минут<sup>3</sup>, а респондент выбирал максимальную цену, которую он готов заплатить за платную поездку с уверенностью, что его поездка не будет отличаться по времени от плана более, чем на 1 минуту. Полученные данные показали, что половина респондентов не готова платить за надёжность времени в пути в принципе (рис. 1). Для другой половины респондентов, которые готовы платить за надёжность времени в пути, было рассчитано средневзвешенное значение стоимости надёжности времени в пути. Оно составило 74,2 рубля.

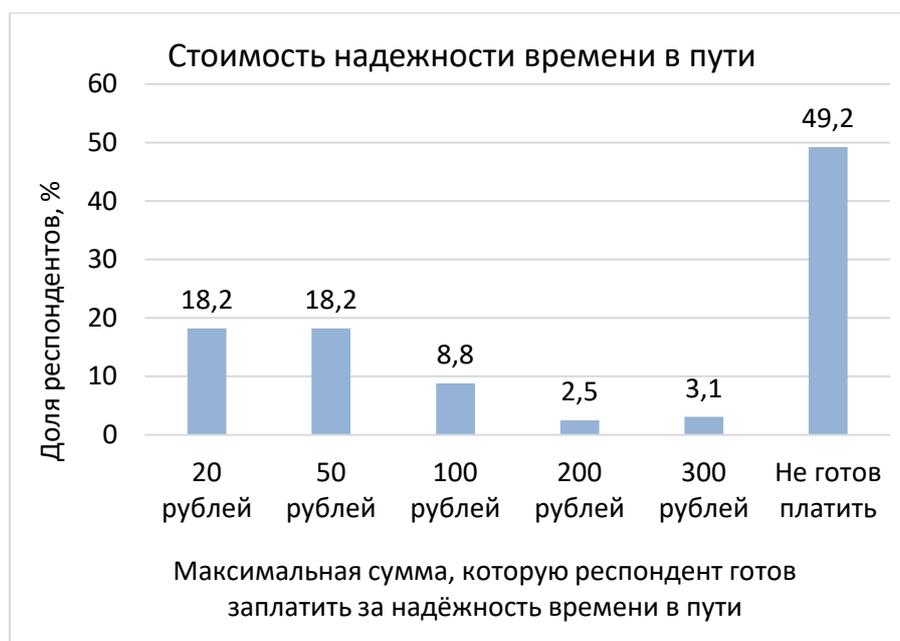


Рис. 1. Стоимость надёжности времени в пути для автомобилистов Владивостока

Таким образом, если платная инфраструктура позволяет не только сэкономить время в пути, но и снизить риски опоздания, часть пользователей готова платить не только за экономию времени как таковую, а ещё и за надёжность, при этом для владивостокцев стоимость надёжности оказалась равна 9 минутам сэкономленного времени. Для того, чтобы отделить друг от друга факторы готовности платить за сэкономленные минуты и за надёжность времени в пути, необходимо проводить дополнительные социологические исследования.

<sup>3</sup> Значения были выбраны исходя из прогнозируемого времени поездки по городу между популярными местами с учётом и без учёта пробок

Тот же опрос показал, что стоимость поездки и её длительность (то есть финансовые и временные затраты) являются довольно важными факторами при принятии решений о поездке, однако как гораздо более важный фактор жители Владивостока указали комфорт в дороге (рис. 2). Таким образом, на готовность платить за поездку влияет также фактор комфорта, и, вероятно, влияние его для жителей разных городов будет разным.

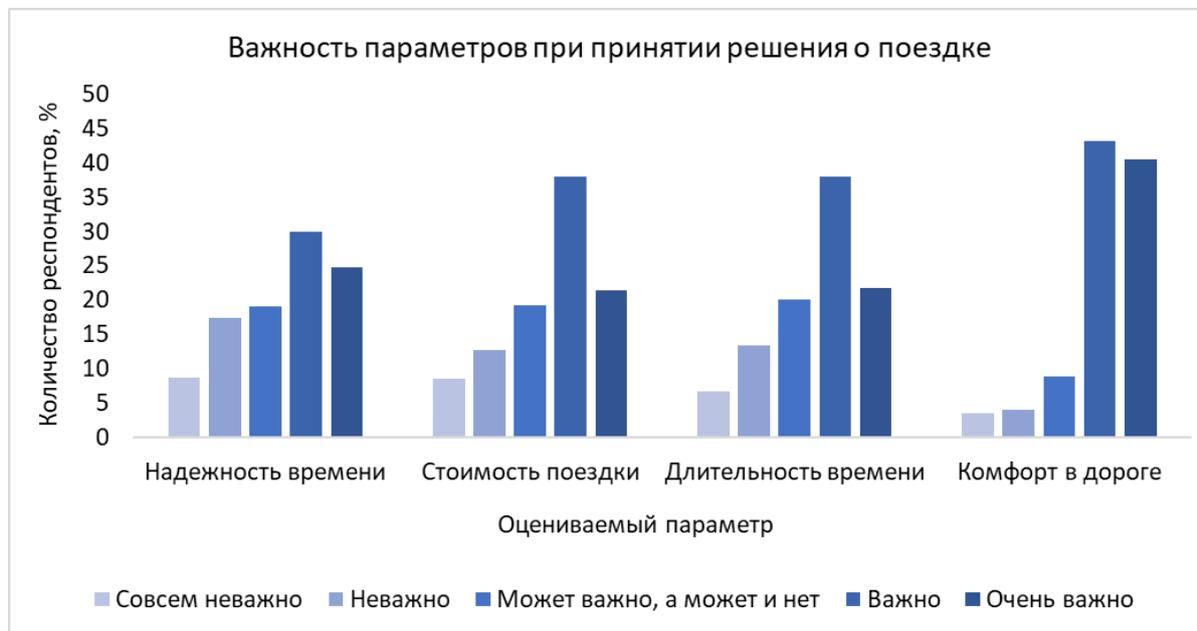


Рис. 2. Важность параметров при принятии решении о поездке для автомобилистов Владивостока

В социологических опросах оценка стоимости экономии времени в пути оценивается обычно на основе ответов респондентов о готовности заплатить за экономию времени в конкретной поездке. Однако готовность платить за такую экономию на регулярной основе, например, ежедневно, может оказаться ниже. Респондентам из Владивостока предлагалось оценить, какой бюджет в месяц они готовы заплатить за экономию в поездках 20 минут каждый день. Для получения стоимости экономии времени в пути на регулярной основе было посчитано и переведено в руб./мин. средневзвешенное значение. Оно составило 3,5 руб./минуту, что значительно ниже показателя 8,2 руб./минуту, полученного из того же опроса, но с учётом ответов респондентов о конкретной поездке.

Таким образом, можно сделать следующие выводы:

- считать среднюю стоимость экономии времени в пути для всех жителей некорректно. Необходимо учитывать социально-демографические характеристики, а также параметры конкретной поездки,
- с увеличением сэкономленного времени в пути стоимость сэкономленного времени растёт нелинейно или не растёт вообще,
- в стоимость экономии времени в пути необходимо включать другие факторы, такие как стоимость надёжности времени в пути, стоимость комфорта, отсутствие пробок и другие. Для каждого конкретного города эти параметры могут различаться,
- стоимость экономии времени в пути отличается при разовой поездке и при регулярных тратах.

Все описанные выводы необходимо учитывать при расчёте стоимости экономии времени в пути для жителей конкретного города. Это позволит более точно прогнозировать тарифы и спрос на платную транспортную инфраструктуру.

### Литература

1. Axhausen Kay & König, Arnd & Abay, Georg & Bates, John & Bierlaire, Michel. Swiss Value of Travel Time Savings, 2006.
2. Mackie P.J. & Jara-Díaz, Sergio & Fowkes, Anthony. The Value of Travel Time Savings in Evaluation // Transportation Research Part E: Logistics and Transportation Review, No. 37, 2001. pp. 91-106.
3. Athiraa I.C., Muneerab C.P., Krishnamurthy K., Anjaneyulud M.V.L.R. Estimation of Value of Travel Time for Work Trips // Transportation Research Procedia, No. 17, 2017. pp. 116-123.
4. Ambarwati L., Indriastuti A.K., Kusumawardhani P. Estimating the Value of Time and Its Application // Open Science Journal, Vol. 2, No. 1, June 2017.
5. Shao C., Liu Y., Liu X. Valuation of Travel Time Savings in Viewpoint of WTA // Computational Intelligence and Neuroscience, 2014.



**Сома Гедеш Мануэл**

Год рождения: 1995

Университет ИТМО,

факультет инфокоммуникационных технологий,

аспирант группы № 7725,

направление подготовки: 09.06.01 – Информатика и

вычислительная техника,

e-mail: guedes.soma@mail.ru

УДК 004.02

**АНАЛИЗ СУЩЕСТВУЮЩИХ СРЕДСТВ  
БИМОДАЛЬНОГО ОПРЕДЕЛЕНИЯ ЭМОЦИЙ  
В ХОДЕ ЗАНЯТИЙ В ДИСТАНЦИОННОМ ФОРМАТЕ**

**Сома Гедеш Мануэл**

**Научный руководитель – д.т.н., доц. О.О. Басов**

**Аннотация**

В эпоху пандемии коронавирусной инфекции большинство стран столкнулось с проблемой адаптации к совершенно новой среде, связанной с переходом на цифровой формат обучения. В работе изучены средства бимодального распознавания эмоций на основе аудио- и видеопотока онлайн-занятий с использованием глубоких сверточных нейронных сетей (CNN), а также семантического слияния аудиовизуальных данных. Цель данной работы – исследовать современные средства бимодального определения эмоций студентов в процессе дистанционного обучения.

**Ключевые слова**

Цифровой формат; средств бимодального распознавания эмоций; аудио данные; визуальные данные; сверточная нейронная сеть (CNN); семантическое слияние.

UDC 004.02

**ANALYSIS OF EXISTING MEANS FOR BIMODAL DETERMINATION  
OF EMOTIONS DURING CLASSES IN REMOTE FORMAT**

**Soma Guedes Manuel**

**Supervisor – Doctor of Technical Sciences, prof. O. Basov**

**Abstract**

In the era of the coronavirus pandemic, most countries have faced the challenge of adapting to a completely new environment associated with the transition to an online digital format. In this work, bimodal emotion recognition tools based on audio and video streams using deep convolutional neural networks (CNNs), as well as semantic fusion of audiovisual data were studied. The purpose of this work is to investigate modern means of bimodal determination of students' emotions in the process of distance learning.

**Keywords**

Digital online format; bimodal emotion recognition means; audio data; visual data; convolutional neural network (CNN); semantic fusion of audiovisual data.

### **Введение**

Дистанционное (удаленное) образование – это форма получения знаний на расстоянии с сохранением компонент учебного процесса и с использованием компьютеров и телекоммуникационных технологий. Данная форма обучения реализуется с преобладанием дистанционных образовательных технологий, методов, форм и средств. Исследования показали [1], что специфика дистанционной формы обучения предъявляет определенные требования к навыкам планирования, принятия решений, а также внимания, памяти и мышления студентов.

Дистанционное обучение также влияет на психоэмоциональную среду учащихся [1]. Психоэмоциональное состояние – это особая форма психических состояний человека с преобладанием эмоционального реагирования по типу доминанты [1]. Иначе говоря, это эмоциональное реагирование человека на какое-то действие, ситуацию или реакцию человека.

Одна из проблем, стоящих перед исследователями взаимодействия человека с компьютером - создание систем, способных распознавать, интерпретировать и реагировать разумным и чувствительным образом на эмоции пользователя [2]. Одним из способов развития данного направления является разработка бимодального средства, определяющего эмоции студентов во время занятий в онлайн-формате. Бимодальность - особенность любого распределения, имеющего две точки или значения - то есть, распределение с двумя модами. В нашем случае - аудио и видео данные. Данное средство позволяет облегчить взаимодействие преподавателя с учеником на основе предоставления информации о психоэмоциональном состоянии ученика во время занятий. На основе аудиовизуальных данных система предоставляет график преподавателю за счет эмоций студента классифицированным во время урока. Несмотря на доступность данных для исследования невербального выражения эмоций, функционирование существующих сегодня средств программного обеспечения ограничено в связи с рядом факторов.

К объективным факторам относят низкое качество предоставляемых каналов связи (причиной чему служит, например, низкая скорость передачи данных) и неподходящая обстановка, в которой используется техническое средство коммуникации. В качестве субъективных факторов выделяется необходимость повышения точности интерпретации эмоций на основании аудиального (голос) и визуального (выражение лица) каналов.

Для повышения эффективности работы с невербальными эмоциями в соответствии с факторами, упомянутыми выше, возникает необходимость разработки средств математического, алгоритмического, методического и программного обеспечения, которые обеспечат повышение эффективности невербальной коммуникации за счет высококачественной интерпретации и учета невербальной составляющей межличностного общения.

### **Сверточная нейронная сеть (CNN)**

Одним из классов нейронных сетей являются CNN. Сверточные сети – это прогресс в глубоком обучении. В девяностые годы CNN использовались для распознавания символов, но они получили более широкое распространение после работы, в которой CNN использовались для классификации изображений. В настоящее время сверточные сети - очень удобный инструмент для тех, кто занимается машинным обучением [3-5].

В работе [6] предложена и оценена эффективная методика распознавания аудиовизуальных эмоций, основанная на применении обученных глубоких сверточных нейронных сетей (CNN). Методика включает в себя 3 основных этапа:

1. Представление необработанных аудиоданных в спектре. В качестве аудиосети выступает предварительно обученная CNN.

2. Обнаружение лиц, обрезка изображений лиц из видео и преобразование лицевых частей в Census-Transformed (CT) в качестве входных данных для предварительно обученной визуальной сети.

3. Извлечение и объединение аудиальных и визуальных признаков из гибридной глубокой сверточной нейронной сети.

### Аудио данные

Чтобы представить необработанные аудиоданные, необходима спектрограмма. Звуки могут быть представлены в форме волн. В данном графике показан пример звуковых волн, имеющих два важных свойства: частоту и амплитуду. Частота определяет высоту звука, амплитуда его громкость.

Далее звуковые волны представлены в спектрограмме, где горизонтальное направление является временем, а вертикальное направлением частотами, изображенной на рис. 1.

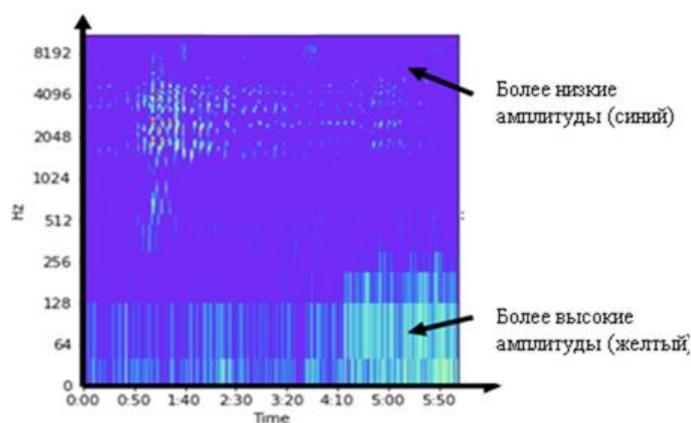


Рис. 1. D спектрограмма [2]

### Визуальные данные

Авторы проводили автоматическое определение лиц с использованием библиотеки Dlib. Извлекались координаты глаз и обрезали верхние и нижние области выражения лица, смотреть рис. 2.



Рис. 2. Обрезанные выражения лиц из набора данных RML. Эмоции (слева направо): гнев, отвращение, страх, счастье, печаль и удивление. (a) обрезанные образцы кадров видео; (б) верхние области выражения лица; (с) нижние области выражения лица [7]

### Сверточная нейронная сеть для обработки аудиовизуальных данных

Метод аудиовизуального распознавания эмоций был представлен с помощью гибридной сверточной нейронной сети (CNN) в работе [8]. Данная архитектура содержит пять входных потоков, авторы сделали эксперименты - с помощью (CNN) извлекали из видеоролика: аудио данные – представлены в мел спектрограмм и визуальные данные - два изображения верхней и два нижней части лица. Далее в схеме показано последовательные блоки, где аудио и визуальные данные объединяются с помощью методов pooling для снижения размеров данных. на этапе Fusing features – объединяет все признаки в один специальный блок. В конце проводится векторизация и подается слой полносвязи нейронных сетей, которые отдельно тренируются для классификации, изображенной на рис. 3.

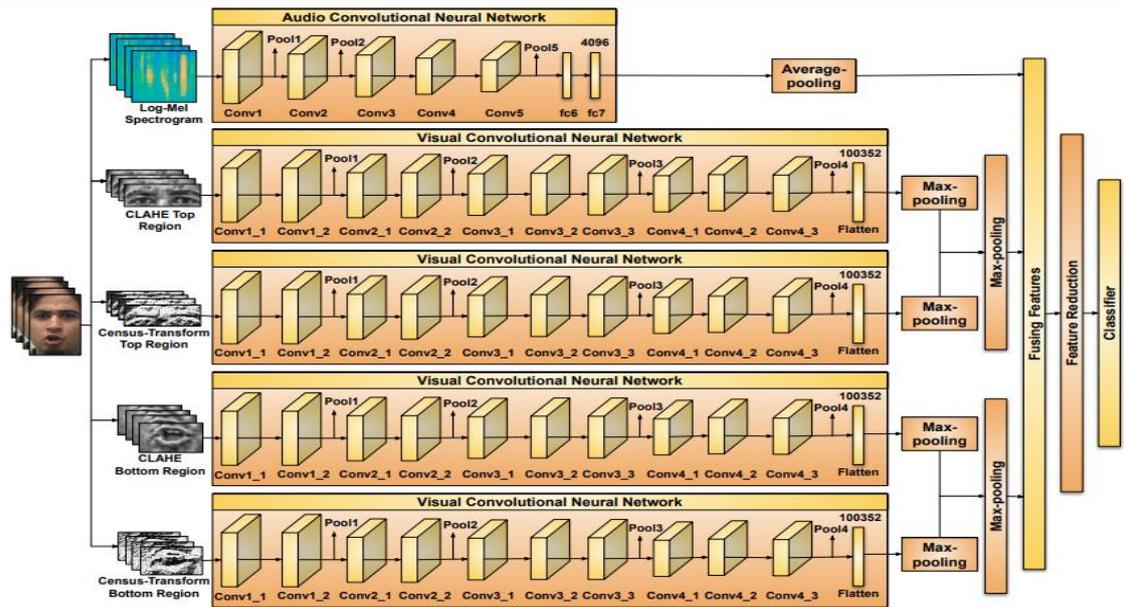


Рис. 3. Схема предлагаемой методики аудиовизуального распознавания эмоций [8]

В таблицах 1 и 2 представлены средние точности (ассигасу) распознавания эмоций в режиме одной модальности. Эти результаты были получены путем вычисления звуковых и визуальных характеристик с помощью предварительно обученной AlexNet и настроенной CNN на основе лицевой части.

Таблица 1

Унимодальность аудиопотока средняя точность (%) [6]

| Набор данных | Метод   | K-NN  | SVM   | LR    | GNB   |
|--------------|---------|-------|-------|-------|-------|
| RML          | PCA     | 37.50 | 53.75 | 60.00 | 43.75 |
|              | PCA+LDA | 66.25 | 65.00 | 64.38 | 63.75 |
| eNTERFACE05  | PCA     | 30.67 | 45.33 | 57.33 | 36.67 |
|              | PCA+LDA | 54.67 | 54.67 | 55.33 | 62.00 |
| BAUM-1s      | PCA     | 31.56 | 37.75 | 40.26 | 38.64 |
|              | PCA+LDA | 33.94 | 45.73 | 40.80 | 40.76 |

Унимодальность видеопотока средняя точность (%) [6]

| Набор данных | Метод   | K-NN  | SVM   | LR    | GNB   |
|--------------|---------|-------|-------|-------|-------|
| RML          | PCA     | 35.00 | 16.25 | 75.00 | 41.25 |
|              | PCA+LDA | 63.75 | 47.50 | 73.75 | 65.00 |
| eNTERFACE05  | PCA     | 44.17 | 30.00 | 65.00 | 34.17 |
|              | PCA+LDA | 67.50 | 60.83 | 66.67 | 65.00 |
| BAUM-1s      | PCA     | 42.21 | 35.36 | 57.95 | 23.22 |
|              | PCA+LDA | 56.77 | 53.19 | 59.43 | 57.14 |

### Семантическое слияния аудиовизуальных данных

В этом разделе мы фокусируемся на распознавании шести прототипных эмоций. Результаты распознавания выражения лица и распознавания эмоций по речи комбинируются с использованием бимодальной мультимодальной модели слияния семантических данных, которая определяет наиболее вероятную эмоцию субъекта. Для распознавания выражения лица используются два типа моделей на основе геометрических черт лица, в зависимости от наличия или отсутствия речи [9].

На рис. 4 показан снимок программной реализации для бимодальной системы распознавания человеческих эмоций. Для обнаружения лиц авторы в основном использовали реализацию метода Viola & Jones из библиотеки OpenCV. Создали процедуры извлечения черт лица, систему распознавания выражений лица и извлечение эмоций из речи на языке программирования C++ [6].



Рис. 4. Снимок программной реализации бимодального алгоритма распознавания эмоций [6]

### Заключение

В данной обзорной работе изучены методы и алгоритмы, классифицирующие человеческие эмоции. На основе лицевой части проведены эксперименты в унимодальных и бимодальных сценариях с использованием методов PCA + LDA, а также несколько алгоритмов классификатора: K-NN, SVM, LR и GNB. В большинстве

случаев классификатор LR достигает более высоких показателей точности (64,38% в аудиальных потоках и 75,00% в визуальных потоках). При сравнительном исследовании можно сделать вывод о том, что метод CNN обладает высокой скоростью работы по распознаванию эмоций, легкостью освоения и обладает более высоким показателем эффективности по сравнению с семантическим слиянием аудиовизуальных данных.

### Литература

1. Соколов Г.А. Особенности психоэмоциональных состояний студента при дистанционной форме обучения. // Современное образование. 2014. № 1. С. 1 – 13.
2. Leandro Yukio Mano Alves. FlexPersuade – Explorando uma abordagem flexível em softwares de persuasão: um estudo de caso com players de música: магистерская диссертация. ком. науки и выч. мат. 2016. С. 81.
3. Yann LeCun Leon Bottou Y.B. Gradient-based learning applied to document recognition / Yoshua Bengio Yann LeCun, Leon Bottou, Patrick Haffner // IEEE. 1998.
4. Лешуков Л.И. Нейронная сеть для оценки музыкальных композиции на основе пользовательских предпочтений. 2016. С. 67.
5. Molau M. Computing mel-frequency cepstral coefficients on the power spectrum / S. Molau, M. Pitz, R. Schluter, H. Ney 2014.
6. Cornejo J. and Pedrini H. Audio-Visual Emotion Recognition using a Hybrid Deep Convolutional Neural Network based on Census Transform// IEEE International Conference on Systems Man and Cybernetics. 2019. С. 1–7.
7. King D.E. Dlib-ml: A Machine Learning Toolkit//Journal of Machine Learning Research. 2009. vol. 10. С. 1755–1758.
8. Learning Affective Features With a Hybrid Deep Model for Audio–Visual Emotion Recognition/ Zhang S., Zhang S., Huang T., Gao W., Tian Q. // IEEE Transactions on Circuits and Systems for Video Technology. 2018. Vol. 2. С. 3030 – 3043.
9. Dragos Datcu, Leon J.M. Rothkrantz. Semantic Audio-Visual Data Fusion for Automatic Emotion Recognition// Man-Machine Interaction. 2015. № 16. С.



**Ставинова Елизавета Алексеевна**

Год рождения: 1996

Университет ИТМО,

факультет цифровых трансформаций,

аспирант,

направление подготовки: 09.06.01 – Информатика

и вычислительная техника,

e-mail: estavinova@icloud.com



**Чунаев Петр Владимирович**

Год рождения: 1987

Университет ИТМО,

факультет цифровых трансформаций,

к.ф.-м.н.,

e-mail: chunaev@itmo.ru

УДК 004.942

## **КЛАССИФИКАЦИЯ КЛИЕНТОВ БАНКА СОГЛАСНО ИНДИВИДУАЛЬНОЙ ПРЕДСКАЗУЕМОСТИ ИХ ПОВЕДЕНИЯ**

**Е.А. Ставинова, П.В. Чунаев**

**Научный руководитель – к.т.н. К.О. Боченина**

Работа выполнена в рамках темы НИР №620188 «Генерация и стратификация данных для оценки и повышения качества предсказательных моделей сложных социэкономических и социотехнических систем».

### **Аннотация**

Цель данной работы заключается в оценке степени предсказуемости события (конкретно, зарубежной поездки клиента банка). Кроме того, в работе предложен метод определения класса предсказуемости клиента, основанный на истории его транзакций. Такой подход позволяет оценить степень предсказуемости зарубежной поездки клиента в следующем месяце до момента ее прогнозирования. Эксперименты показывают, что данный подход достаточно эффективен.

### **Ключевые слова**

Предсказуемость, транзакционные данные, LSTM, машинное обучение, нейронная сеть.

### **Введение**

Задача прогнозирования поведения клиента является чрезвычайно популярной в настоящее время и имеет множество применений от прогнозирования следующей покупки до определения будущего местоположения человека [1, 2]. В большинстве подходов прогнозы получаются путем обучения моделей на части данных, а качество предсказания измеряется на тестовых данных с помощью некоторых метрик. Целью данной работы является анализ степени предсказуемости события (точнее, зарубежной поездки) до момента прогнозирования. Степень предсказуемости может дать ответы на вопросы: стоит ли делать прогноз события для этого клиента? Есть ли смысл в получении ответов модели в данном случае?

Основной идеей данной статьи является предположение о том, что массив клиентов может быть разделен на классы в зависимости от степени предсказуемости зарубежной поездки в следующем месяце согласно истории транзакций клиентов. После разделения клиентов на классы двунаправленная сеть с долгой краткосрочной памятью (BiLSTM) обучается на некотором наборе данных для определения класса предсказуемости клиента до момента прогнозирования. Обученная сеть может быть использована в дальнейшем для определения класса предсказуемости новых клиентов банка или для переопределения класса клиента в случае изменения его поведения с точки зрения путешествий. Знание класса предсказуемости клиента дает возможность оценить качество прогнозирования поездки для данного клиента до момента прогнозирования.

Существующие работы в области предсказуемости посвящены измерению предсказуемости временных рядов и признаков [3, 4]. Следует отметить, что целью анализа предсказуемости временных рядов является оценка возможности существования некоторых паттернов во временном ряде, в то время как анализ предсказуемости признаков направлен на выделение набора таких признаков, которые будут полезны для прогнозной модели. Новизна данного исследования заключается в создании метода оценки предсказуемости события и использовании данного метода для определения эффективности модели до момента прогнозирования. Под термином предсказуемости события в данном случае принимается предсказуемость временного ряда, состоящего из категориальных значений, которые являются бинарными индикаторами события. В данной работе для прогнозирования событий была использована нейронная сеть с долгой краткосрочной памятью (LSTM), так как эта разновидность архитектуры имеет преимущество в усвоении временных зависимостей. Ранее данная сеть использовалась для прогноза местоположения клиента, и результаты экспериментов доказывают ее применимость к этой задаче [5].

Актуальность предложенного в работе метода прогнозирования и оценки уровня предсказуемости обусловлена использованием исключительно исторических данных о транзакциях клиента, совершенных с помощью дебетовых карт. Поэтому данный подход может быть использован в качестве основы для разработки выгодных ставок для тех клиентов, чье поведение с точки зрения путешествий может быть предсказано моделью с достаточным уровнем качества. Кроме того, с применением предложенного подхода появится возможность стратификации клиентов в соответствии с их уровнем предсказуемости и создания персонализированных рекламных кампаний согласно их уровню предсказуемости.

### **Методология исследования**

#### **Метод оценки предсказуемости**

В данной работе рассматривается задача прогнозирования событий и оценки их предсказуемости, в связи с чем было разработано следующее определение предсказуемости события. Если функция распределения значений метки события  $label$  в момент  $t$  не зависит от ее предыдущих значений, т.е.:  $F(label_t | \Omega_{t-1}) = F(label_t)$ , где  $F$  – функция распределения,  $\Omega_{t-1}$  – предыдущие значения метки события, то метка события является непредсказуемой. В определении под термином метки события принимается бинарный индикатор события в момент времени  $t$  (значение 1 показывает, что событие имеет место, 0 – обратную ситуацию).

Что касается подхода к оценке степени предсказуемости, то одним из возможных вариантов является подход, основанный на измерении ошибки прогнозирования. В таком случае степень предсказуемости события является по сути выборочной оценкой предсказуемости, поскольку ее вычисление связано с измерением ошибок прогноза на имеющихся данных. В данном исследовании была использована следующая

выборочная оценка предсказуемости с достаточно простой формулой:

$$C(n, Q, m) = 1 - \frac{1}{n} \sum_{i=1}^n |Q_i - \widehat{Q}_i| \in [0,1],$$

где  $n$  – длина тестового периода,  $Q_i$  – реальное значение индикатора события,  $\widehat{Q}_i$  – предсказанная вероятность события,  $m$  – модель, которая использовалась для прогнозирования.

Предложенный коэффициент  $C$  используется для разделения массива клиентов на классы в соответствии со степенью предсказуемостью факта зарубежной поездки в следующем месяце для них. Очевидно, что данное деление может быть произведено только после того, как будут получены прогнозы модели. Далее в работе будут описаны два класса предсказуемости, выделенные исходя из значений коэффициента. Данное разбиение было мотивировано желанием отделить объекты с предсказуемым поведением от объектов, чье поведение непредсказуемо.

### Метод определения класса предсказуемости клиента

Класс предсказуемости клиента банка может быть установлен путем получения прогнозов модели и вычисления значения коэффициента  $C$ . Таким образом, для определения класса предсказуемости необходимо знать значение коэффициента предсказуемости, которое должно быть получено посредством сравнения прогнозов модели с реальными данными. После этого становится возможным оценить, насколько предсказуемо поведение того или иного клиента. Однако, на практике было бы чрезвычайно полезно иметь возможность пропустить шаг с использованием прогнозной модели и делать вывод о предсказуемости поведения конкретного клиента с точки зрения путешествий, используя исключительно последовательность индикаторов событий, соответствующих нескольким последним месяцам. В связи с этим в рамках данной работы был создан алгоритм определения класса предсказуемости по последовательности индикаторов событий.

Фактически данный алгоритм является одним из способов решения задачи классификации. Цель задачи состоит в том, чтобы определить, к какому классу предсказуемости относится клиент банка, описанный только векторами признаков тренировочного периода. Такая формулировка задачи максимально приближена к действительности, поскольку в определенный момент времени доступ может быть предоставлен лишь к данным периода наблюдений без возможности запроса данных будущего (т.е. данных из тестового периода).

Для решения поставленной задачи классификации последовательностей используется двунаправленная сеть с долгой краткосрочной памятью (BiLSTM). На вход сети поступает множество последовательностей с категориальными значениями, которые являются индикаторами события; временной шаг между элементами последовательности – один месяц. Длина последовательности была выбрана равной шести. На выходе сеть выдает принадлежность классу предсказуемости для каждой последовательности, поданной на вход. Для обучения данной модели были извлечены последовательности, описывающие поведение клиентов с точки зрения зарубежных поездок в течение последних шести месяцев тренировочного периода, а затем был установлен их класс предсказуемости путем использования модели прогнозирования, ее ответов и коэффициента  $C$ . После обучения классификационной модели на последовательностях за последние шесть месяцев тренировочного периода было измерено качество ее работы с помощью следующих шести месяцев истории транзакций. Основная идея метода определения класса предсказуемости клиента заключается в получении обученной модели, которая может быть использована в дальнейшем в случае поступления новых данных: для переопределения класса предсказуемости клиента или для идентификации класса предсказуемости нового клиента банка.

## Результаты экспериментов

### Данные и их предобработка

Для экспериментов, проведенных в рамках исследования, был использован набор данных, предоставленный одним российским банком. Эти данные не находятся в общем доступе из-за политики банка. Набор данных состоит из транзакций, совершенных клиентами (6 тыс.) в 2018 году (с января по декабрь).

Первым этапом обработки данных является категоризация транзакций с помощью кода категории продавца (МСС). Все коды были разделены на 87 категорий. Затем были сформированы две группы транзакций по месту совершения: российские и зарубежные. Эта процедура была проведена с использованием ISO-кода страны, указанного в деталях каждой транзакции. Для формализации поведения клиента были использованы векторы признаков, характеризующие его расходы в определенном месяце. Эти векторы содержат информацию о сумме (в рублях), потраченной клиентом в данном месяце в каждой категории. После этого к векторам признаков было применено преобразование Йео-Джонсона. В полученных 88-мерных векторах первые 87 координат соответствуют 87 категориям МСС кодов, а 88-я координата – метка локации клиента в конкретном месяце. Целевой переменной является метка локации клиента в следующем месяце. В данном исследовании было решено присваивать метку зарубежной локации месяцу, в течение которого клиент совершил хотя бы одну зарубежную транзакцию. Следует отметить, что транзакции, совершенные в сети Интернет, были помечены как российские.

После этапа предобработки данных и составления векторов признаков все полученные векторы были разделены на подпоследовательности следующим образом. Каждая подпоследовательность содержит пять векторов (т.е. месяцев); все векторы одной последовательности описывают траты одного клиента; последовательность непрерывна в течение этих пяти месяцев. Далее были выделены тренировочный (с января по июнь) и тестовый (с июля по декабрь) периоды. Все подпоследовательности, заканчивающиеся месяцами из тестового периода, были отведены для проверки качества работы модели, которая обучалась на векторах признаков из тренировочного периода. Пример разделения векторов признаков на периоды представлен на рис. 1.

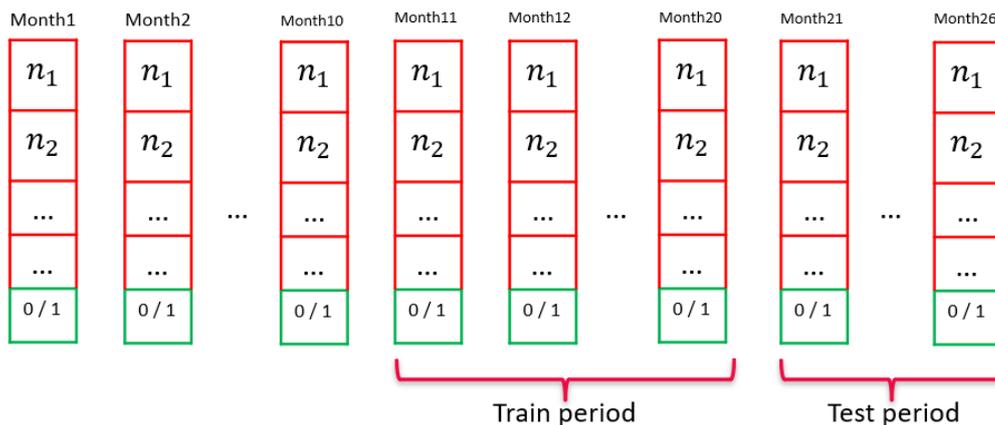


Рис. 1. Пример разделения множества векторов признаков на тренировочный и тестовый периоды

Для прогнозирования факта совершения зарубежной поездки клиентами банка использовалась сеть LSTM. Количество скрытых слоев и количество нейронов в них было определено с помощью нескольких экспериментов с различными настройками. Было проведено шесть экспериментов с различными конфигурациями сети: 1, 2, 3 скрытых слоя и 32, 64 нейрона в каждом скрытом слое. По результатам экспериментов

была выбрана комбинация из двух скрытых слоев с 64 нейронами. Значение коэффициента скорости обучения было принято 0,001, основываясь на очередной серии экспериментов.

В текущей постановке задачи целью является нахождение ответа на вопрос: насколько предсказуем факт совершения зарубежной поездки клиентом банка в следующем месяце? Или, точнее, насколько предсказуем факт совершения зарубежной поездки клиентом банка в следующем месяце при использовании описанной модели LSTM с векторами признаков в качестве входных данных?

### Оценка предсказуемости

Ранее в статье была введена мера для оценки степени предсказуемости, названная коэффициент предсказуемости  $C$ . В основе коэффициента лежат прогнозы модели, которая используется для предсказания факта наступления события (зарубежной поездки). Значения коэффициента  $C$  были вычислены для каждого клиента банка, представленного в наборе данных, с помощью следующей процедуры: (а) модель, основанная на сети LSTM, обучалась на векторах признаков тренировочного периода; (б) после получения ответов модели для каждого клиента рассчитывалось значение коэффициента с использованием действительных индикаторов события, принадлежащего тестовому периоду, спрогнозированных моделью вероятностей события и длины тестового периода. На рис. 2 показано распределение значений коэффициента  $C$ .

Следующим этапом работы являлось создание классов предсказуемости на основе значений коэффициента предсказуемости. Было решено разделить клиентов на классы путем сравнения значения  $C$ , соответствующего данному клиенту, с медианным значением  $C$ . Такое деление обеспечивает классы одинакового размера, причем параметр деления не должен настраиваться индивидуально для разных выборок. В данном случае медианное значение коэффициента равно 0,68. Далее в соответствии со значением коэффициента были сформированы два класса: низкой и высокой предсказуемости. Распределение клиентов банка по классам предсказуемости представлено в табл. 1. Клиенты класса высокой предсказуемости имеют более предсказуемое поведение с точки зрения модели LSTM, поскольку для них ответы модели ближе к реальности по сравнению с ответами для клиентов класса низкой предсказуемости.

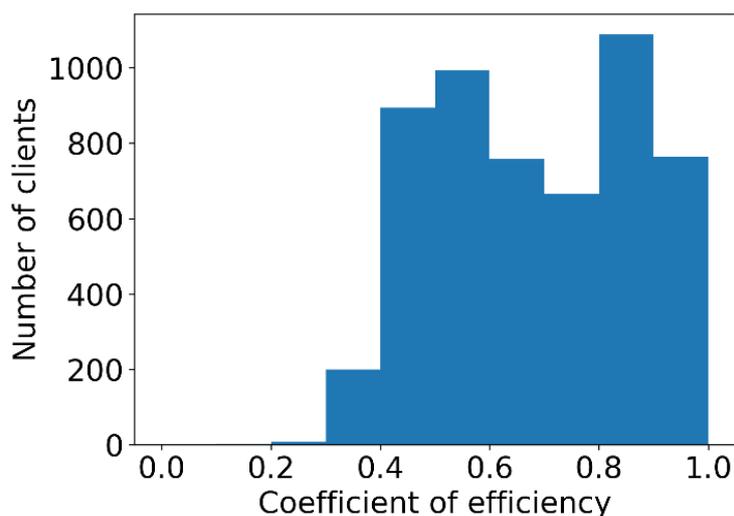


Рис. 2. Распределение значений коэффициента  $C$

### Распределение клиентов банка по классам предсказуемости

| Класс предсказуемости | Значения коэффициента   | Количество представителей |
|-----------------------|-------------------------|---------------------------|
| Высокой               | $[\text{median}(C); 1]$ | 2687                      |
| Низкой                | $[0; \text{median}(C))$ | 2687                      |

После выделения классов предсказуемости было проведено исследование качества прогноза для каждого класса. Для оценки качества были использованы метрики Precision и Recall. На рис. 3 показаны Precision-Recall кривые для классов предсказуемости, описанных в табл. 1. Кроме того, на рис. 3 изображена прямая, соответствующая частоте совершения зарубежных поездок в тестовом наборе данных (0,51), которая позволяет сравнивать качество классификаторов с качеством случайного угадывания. Для изученного набора данных предложенный классификатор показал качество выше, чем у случайного угадывания.

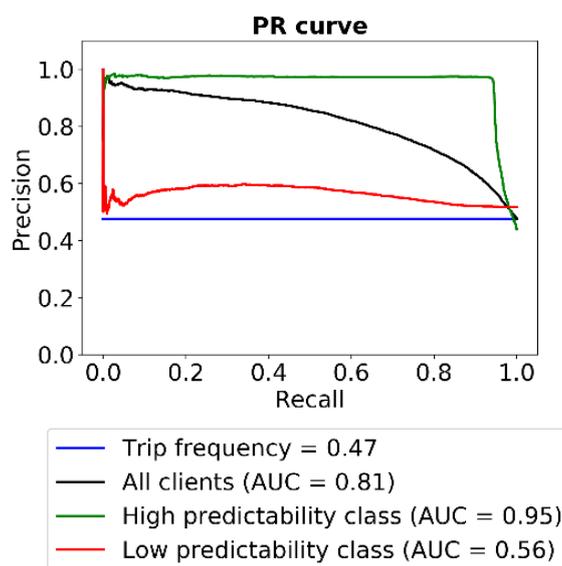


Рис. 3. Precision-Recall кривые для классов предсказуемости, полученных после момента прогнозирования, для прогноза на шесть месяцев вперед

В результате разделения массива клиентов на классы предсказуемости был выявлен класс с почти идеальной Precision-Recall кривой и класс с кривой, расположенной чуть выше, чем прямая, соответствующая точности случайного угадывания. Это означает, что предлагаемое разделение массива клиентов на классы предсказуемости является полезным.

### Определение класса предсказуемости клиента

Выше был введен метод разделения массива клиентов на два класса предсказуемости с использованием коэффициента  $C$ . Данный метод предполагает, что прогнозы модели уже получены и значения коэффициента вычисляются после момента прогнозирования. В этом разделе будет сделан упор на задачу определения класса предсказуемости клиента до момента прогнозирования. Поставленная задача может быть сведена к задаче определения класса предсказуемости клиента с помощью данных, относящихся исключительно к тренировочному периоду. Эта задача была

решена с использованием алгоритма, предложенного в разделе Метод определения класса предсказуемости клиента. Матрица ошибок, позволяющая судить о точности определения класса предсказуемости, представлена в табл. 2. Точность определения класса предсказуемости была измерена с помощью метрики Accuracy, значение которой составило 86,32%. Результаты определения класса предсказуемости достаточно сбалансированы и точны.

Таблица 2

**Матрица ошибок определения класса предсказуемости**

|            | Прогноз                       |                               |                              |
|------------|-------------------------------|-------------------------------|------------------------------|
|            |                               | Класс высокой предсказуемости | Класс низкой предсказуемости |
| Реальность | Класс высокой предсказуемости | 2050                          | 544                          |
|            | Класс низкой предсказуемости  | 170                           | 2458                         |

### **Заключение**

В данной работе был предложен подход для оценки предсказуемости события (зарубежной поездки клиента банка) после получения прогнозов модели. Для оценки предсказуемости была предложена специальная мера, в основе которой лежит принцип сравнения реальных индикаторов события и вероятности наступления события, спрогнозированной моделью. После момента прогнозирования массив клиентов банка был разделен на две группы в соответствии с предсказуемостью события для них с использованием значений меры. Затем был предложен метод определения класса предсказуемости клиента на основании исторических данных до момента прогнозирования. Основным достижением этой работы является возможность выделения таких клиентов, поведение которых может быть предсказано с почти идеальным качеством. Зная принадлежность клиента к одному из классов предсказуемости, становится возможной оценка качества прогноза события для этого клиента до момента прогнозирования.

### **Литература**

1. Vaganov D. et al. Forecasting Purchase Categories with Transition Graphs Using Financial and Social Data // *Social Informatics* / ed. Staab S., Koltsova O., Ignatov D.I. Cham: Springer International Publishing, 2018. P. 439–454.
2. Moon G.E., Hamm J. A Large-Scale Study in Predictability of Daily Activities and Places // *MobiCASE'16 Proceedings of the 8th EAI International Conference on Mobile Computing, Applications and Services*. 2016. P. 86–97.
3. Kaboudan M.A. A measure of time series' predictability using genetic programming applied to stock returns // *J. Forecast.* 1999. Vol. 18, № 5. P. 345–357.
4. Javed K. et al. Features selection procedure for prognostics: An approach based on predictability // *IFAC Proceedings Volumes (IFAC-PapersOnline)*. IFAC Secretariat, 2012. Vol. 45, № 20 PART 1. P. 25–30.
5. Stavinova E., Bochenina K. Forecasting of foreign trips by transactional data: A comparative study // *Procedia Computer Science*. Elsevier B.V., 2019. Vol. 156. P. 225–234.



**Тимофеева Елена Павловна**

Год рождения: 1997

Университет ИТМО,

факультет информационных технологий и программирования,  
студент группы №М42212,

направление подготовки: 09.04.02 – Информационные  
системы и технологии,

e-mail: etimofeeva277@gmail.com



**Астапов Сергей Сергеевич**

Год рождения: 1988

Университет ИТМО,

факультет информационных технологий и программирования,  
PhD,

e-mail: sergeiastapov@gmail.com

УДК 004.93

**ОПРЕДЕЛЕНИЕ ПЕРЕКРЫТИЙ  
В РЕЧИ С ПОМОЩЬЮ МОДЕЛИ COUNTNET**

**Е.П. Тимофеева**

**Научный руководитель – PhD С.С. Астапов**

**Аннотация**

В работе проведены исследования по определению перекрытий в речи на записях, где есть речь одного или более дикторов, с помощью метода CountNet. Этот метод был разработан для задачи определения количества дикторов в записи, но в ходе этой работы был адаптирован для задачи определения перекрытий речи в записи. Все эксперименты проводились на наборе данных LibriCount.

**Ключевые слова**

Детектор перекрытой речи, countnet, libricount, сверточная нейронная сеть, рекуррентная нейронная сеть.

Наличие перекрывающейся или совмещенной речи – это обычное явление и естественное следствие спонтанных многосторонних разговоров, возникающих при общении нескольких людей. В человеческой речи есть несколько случаев, когда в одно и то же время говорят несколько людей. Это может быть перебивание речи, короткие ответы на фоне (например, «да», «ага») или простые кратковременные естественные перекрытия во время быстрой смены диктора.

Такая речь представляет собой серьезную проблему для автоматических систем, обрабатывающих эти аудиоданные. Например, в задаче определения диктора в современных системах речевые сегменты назначаются только одному диктору, что приводит к ошибкам пропущенной речи в местах, где активно говорят более одного человека.

В данной работе были проведены исследования по определению перекрытий в речи на записях, где есть речь одного или более дикторов рис. 1. Для решения этой

задачи был выбран метод CountNet [1]. Целью этого метода является определение количества дикторов в записи.

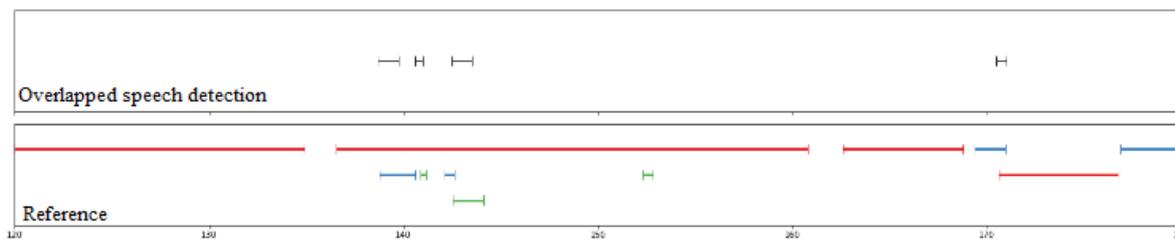


Рис. 1. Определение перекрытий в речи [5]

В методе CountNet представлены несколько моделей: сверточная нейронная сеть (CNN), рекуррентная нейронная сеть (RNN), сверточная рекуррентная нейронная сеть (CRNN) и сверточные рекуррентные нейронные сети с полным диапазоном (F-CRNN) рис. 2 [2].

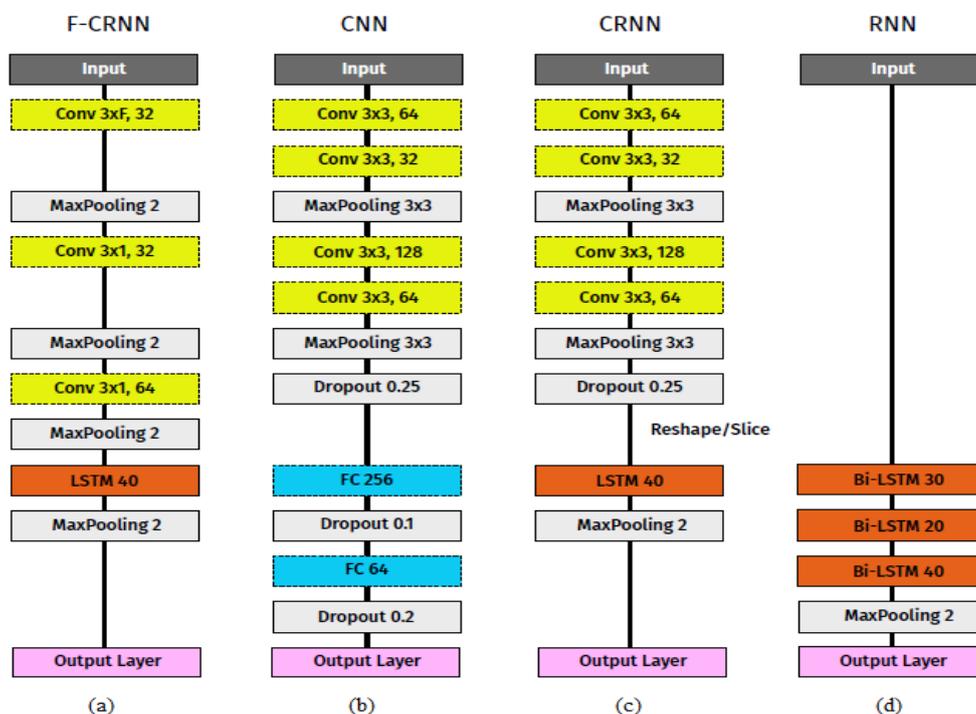


Рис. 2. Архитектуры моделей CountNet [2]

Большую роль в определении дикторов играет продолжительность аудиозаписи. Так как CountNet ориентирована на задачу определения количества дикторов, то авторами был предложен эксперимент для того, чтобы проверить качество работы системы на небольших отрезках записи. Этот эксперимент был поставлен с использованием модели CRNN. Для этого на вход подавались отрезки записей продолжительностью от 1 до 9 секунд рис. 3. И по этим данным можно сделать вывод, что модель показывает хорошие результаты даже на коротких записях. И использование этой системы хорошо подойдет для решения задачи определения перекрытий в речи [4, 5].

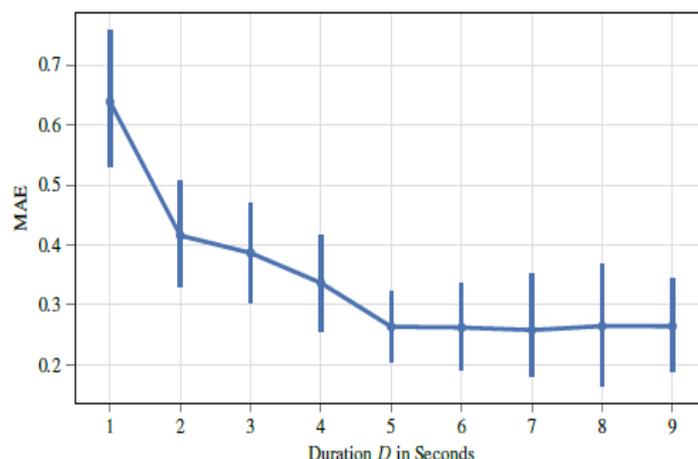


Рис. 3. Результаты эксперимента с различными продолжительностями аудиофайла [2]

Эксперименты проводились на предоставленных предобученных моделях. В таблице приведены результаты работы всех моделей на наборе данных LibriCount [3]. По этим результатам видно, что все модели показали отличный результат. Основной причиной таких результатов является предобученность этих моделей. А также то, что метод CountNet предназначен для определения количества дикторов на записи, и это объясняет, то почему этот метод также хорошо работает и на задаче определения пересечений в речи. Обе эти задачи тесно связаны между собой, и система, которая хорошо себя показала на одной задаче, также подойдет и для другой (таблицы).

Таблица

#### Результаты работы моделей CountNet

|                  | CNN   | RNN   | CRNN  | F-CRNN |
|------------------|-------|-------|-------|--------|
| <i>Mean MAE</i>  | 0.076 | 0.004 | 0.011 | 0.009  |
| <i>Accuracy</i>  | 0.92  | 0.94  | 0.99  | 0.99   |
| <i>Precision</i> | 0.96  | 0.99  | 0.99  | 0.99   |
| <i>Recall</i>    | 0.79  | 0.99  | 0.97  | 0.98   |
| <i>F1-score</i>  | 0.85  | 0.99  | 0.98  | 0.98   |

Модель CountNet имеет множество применений в аудиоанализе, и в частности для задачи определения перекрытий в речи. И исходя из этих результатов, можно сделать вывод, что эта модель хорошо подходит для данной задачи.

В дальнейшей работе планируется применение архитектур CountNet на других наборах данных и сравнение этих результатов с результатами других популярных методов.

#### Литература

1. Astapov S., Popov D., Kabarov V. (2020) Directional Clustering with Polyharmonic Phase Estimation for Enhanced Speaker Localization. In: Karpov A., Potapova R. (eds) Speech and Computer. SPECOM 2020. Lecture Notes in Computer Science, vol 12335. Springer, pp. 45-56.
2. CountNet [электронный ресурс]. Режим доступа: <https://github.com/faroit/CountNet> – (Дата доступа 18.12.2020).

3. Fabian-Robert Stöter, Soumitro Chakrabarty, Bernd Edler, Emanuël Habets. CountNet: Estimating the Number of Concurrent Speakers Using Supervised Learning. *IEEE/ACM Transactions on Audio, Speech and Language Processing*, Institute of Electrical and Electronics Engineers, 2019, *IEEE/ACM Transactions on Audio, Speech, and Language Processing*, 27 (2), pp.268-282. 10.1109/TASLP.2018.2877892. lirmm-02010805.
4. LibriCount, a dataset for speaker count estimation [электронный ресурс]. Режим доступа: <https://zenodo.org/record/1216072#.X-VnDb3VK02>. (Дата доступа 18.12.2020).
5. Neural speaker diarization with pyannote-audio [электронный ресурс]. Режим доступа: <https://github.com/pyannote/pyannote-audio>. (Дата доступа 10.12.2020).



**Тихонова Татьяна Алексеевна**  
Университет ИТМО,  
факультет цифровых трансформаций,  
студент группы №J41011,  
направление подготовки: 09.04.68.02 – Цифровые геотехнологии,  
e-mail: ttikhonova98@gmail.com



**Хватов Александр Александрович**  
Университет ИТМО,  
к.ф.-м.н.,  
с.н.с. национальный центр когнитивных разработок,  
e-mail: alex\_hvatov@itmo.ru

УДК 519.63

## **ЧИСЛЕННОЕ РЕШЕНИЕ ДИФФЕРЕНЦИАЛЬНЫХ УРАВНЕНИЙ В ЧАСТНЫХ ПРОИЗВОДНЫХ С ПОМОЩЬЮ МЕТОДОВ ОПТИМИЗАЦИИ**

**Т.А. Тихонова, А.А. Хватов**

**Научный руководитель – к.ф.-м.н., с.н.с. А.А. Хватов**

Работа выполнена в рамках темы НИР №620175 «Виртуальная лаборатория разработки композитных моделей с помощью эволюционных методов автоматического машинного обучения».

### **Аннотация**

Большинство существующих методов численного решения дифференциальных уравнений в частных производных (ДУЧП) имеют ограничения в применимости. В этой работе предлагается метод, позволяющий решать ДУЧП произвольного вида. Показана сходимость и устойчивость предложенного метода на волновом уравнении и уравнении теплопроводности.

### **Ключевые слова**

Численное решение ДУЧП, дифференциальные уравнения в частных производных, метод оптимизации.

В последнее время появляются методы, позволяющие получать дифференциальные уравнения в частных производных (ДУЧП) по измеренным данным физического процесса [1]. Для последующего решения полученных уравнений с другими граничными и начальными условиями необходимо уметь численно решать ДУЧП произвольного вида, так как вид полученного уравнения заранее неизвестен.

Большинство существующих методов численного решения ДУЧП построены для решения определенных классов задач. Сходимость метода конечных разностей [2] и метода конечных элементов доказана для ограниченного класса уравнений. Нейронные сети [3] тоже имеют ограничения применимости, а также требуют время на обучение для решения новых уравнений. Решение, основанное на спектральных методах, [4], также имеет ограничения на вид дифференциального оператора и граничных условий.

Предполагается, что предложенным в этой работе методом оптимизации можно решать произвольные ДУЧП.

В методе оптимизации на области, на которой ищется решение, строится сетка, и решение ищется в узлах сетки. В узлах сетки производится разностная аппроксимация частных производных как в методе конечных разностей. Для центральных узлов используется следующая схема:

$$u'_x = \frac{u_{x_i+h} - u_{x_i-h}}{2h}. \quad (1)$$

В выражении (1)  $u_{x_i}$  означает значение искомой функции в узле  $x_i$ ,  $h$  - шаг сетки по  $x$ .

Для граничных узлов используются односторонние производные:

$$u'_x = \frac{u_{x_0+h} - u_{x_0}}{h},$$
$$u'_x = \frac{u_{x_N} - u_{x_N-h}}{h}.$$

Чтобы получить производные высших порядков, тот же шаблон применяется несколько раз. Если в граничных условиях есть производная, она аппроксимируется той же схемой.

Решение в узлах ищется с помощью минимизации разности правой и левой частей уравнения вместе с граничными условиями, умноженными на коэффициент  $\lambda$ :

$$\|Lu - f\| + \lambda \|Bu - f_b\| \xrightarrow{u} \min. \quad (2)$$

В выражении (2)  $L$  – дифференциальный оператор, действующий на сетку  $u$ ,  $f$  – правая часть дифференциального уравнения,  $B$  – дифференциальный оператор в граничном условии,  $f_b$  – правая часть граничного условия.

Чтобы показать сходимость и устойчивость метода, методом оптимизации были решены волновое уравнение и уравнение теплопроводности с граничными условиями первого рода.

Для этих уравнений было сделано примерно 30 запусков для разных размеров сеток от  $10 \times 10$  до  $50 \times 50$  со случайным начальным приближением и рассчитаны МАЕ и время работы алгоритма (рис. 1, 2). Результат оптимизации практически не зависит от начального приближения. Средняя абсолютная ошибка падает с увеличением разбиения сетки, что показывает сходимость метода.

Время работы алгоритма сильно увеличивается с увеличением размера сетки и для расчета волнового уравнения на сетке  $50 \times 50$  составляет почти 3 часа. Его можно заметно уменьшить, если вместо случайного начального приближения использовать интерполированное решение задачи на меньшей сетке.

В таблице в первой строчке можно увидеть время расчета сетки  $50 \times 50$  со случайным начальным приближением. Во второй строчке показано время вычисления сетки  $50 \times 50$  с решением  $25 \times 25$  в качестве начального приближения. Суммарное время работы второй строчки меньше, чем время работы со случайным начальным приближением.

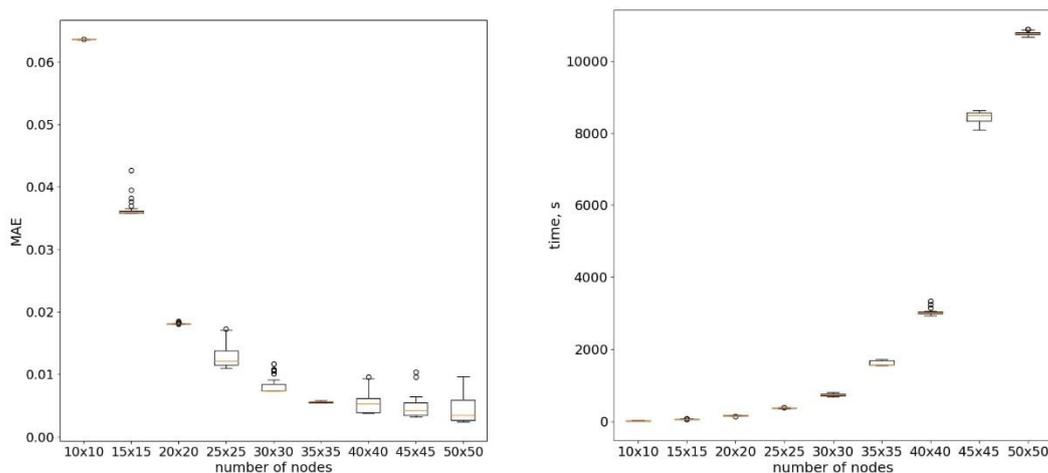


Рис. 1. MAE и время вычисления волнового уравнения для разных размеров сеток

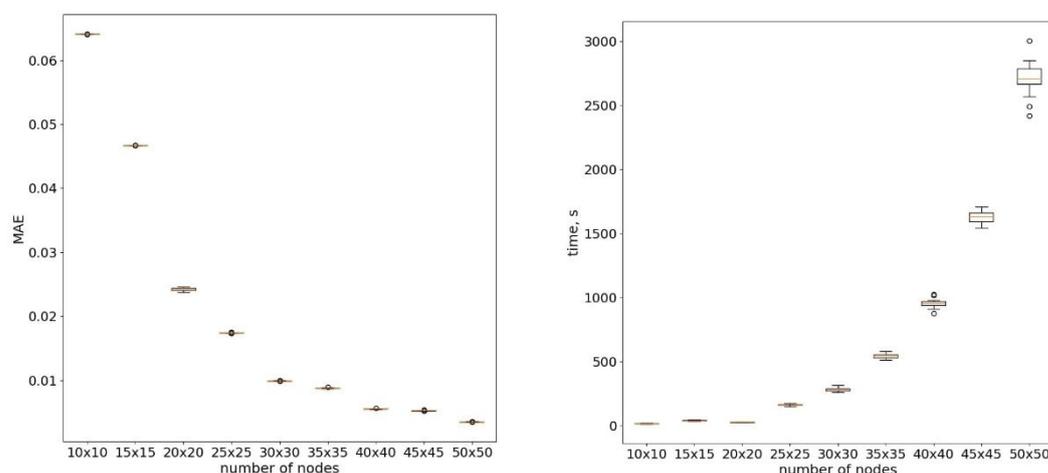


Рис. 2. MAE и время вычисления уравнения теплопроводности для разных размеров сеток

Таблица

**Время вычислений при случайном и заданном начальных приближениях**

| Начальное приближение | 25×25 | 50×50 | Сумма |
|-----------------------|-------|-------|-------|
| Случайное             | -     | 7848  | 7848  |
| Интерполированное     | 156   | 7008  | 7164  |

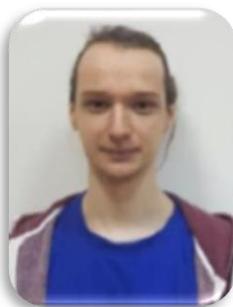
В работе был предложен метод решения произвольных ДУЧП. Проведенные эксперименты с волновым уравнением и уравнением теплопроводности показали сходимость и устойчивость метода для гиперболических и параболических уравнений. Недостатком данного метода является низкая скорость работы. Для уменьшения времени работы алгоритма на больших сетках в качестве начального приближения было использовано решение меньших сеток.

В качестве дальнейшей разработки планируется ввести решение многомерных уравнений, ускорить работу алгоритма, запуская его на GPU, разработать методы

получения лучшего начального приближения, ввести неравномерную сетку, задание непостоянных коэффициентов.

### Литература

1. Maslyaev M., Hvatov A., Kalyuzhnaya A. Data-Driven Partial Differential Equations Discovery Approach for the Noised Multi-dimensional Data //International Conference on Computational Science. Springer, Cham, 2020. С. 86-100.
2. Калиткин Н.Н. Численные методы. 2 изд. БХВ-Петербург. 2011.
3. Li Z. et al. Fourier neural operator for parametric partial differential equations //arXiv preprint arXiv:2010.08895. 2020.
4. Townsend A., Olver S. The automatic solution of partial differential equations using a global spectral method //Journal of Computational Physics. 2015. Т. 299. С. 106-123.



**Хрульков Александр Александрович**  
Университет ИТМО,  
Институт дизайна и урбанистики,  
аспирант группы №7738,  
направление подготовки: 05.13.18 – Математическое  
моделирование, численные методы и комплексы программ,  
e-mail: oneonwar@itmo.ru



**Терентьев Роман Андреевич**  
Университет ИТМО,  
факультет программной инженерии и компьютерной техники,  
магистрант группы №Р41143,  
направление подготовки: 09.04.04 – Программная инженерия,  
e-mail: Jason03091997@yandex.ru



**Митягин Сергей Александрович**  
Университет ИТМО,  
Институт дизайна и урбанистики,  
к.т.н.,  
e-mail: mityagin@itmo.ru

УДК 51-7

**КОМПЬЮТЕРНОЕ ЗРЕНИЕ ДЛЯ АНАЛИЗА  
И ПОДСЧЕТА ГОРОДСКОГО ТРАФИКА  
А.А. Хрульков, Р.А. Терентьев  
Научный руководитель – к.т.н. С.А. Митягин**

Работа выполнена в рамках темы НИР №620181 «Средства определения тематики и значимости объектов городской среды на основе открытых данных социальных сетей».

**Аннотация**

Работа посвящена возможности использования технологий компьютерного зрения для камер городского видеонаблюдения для подсчета городского трафика, и использования этих данных в целях анализа городской среды. Проанализированы существующие методы компьютерного зрения для подсчета трафика. Сформирован подход, позволяющий получать качественные данные о городском трафике, пригодные для использования геоаналитических методов и анализа городской среды.

**Ключевые слова**

Компьютерное зрение, городской трафик, геоинформатика, машинное обучение, городская среда, урбанистика.

Получение качественных данных при анализе городского трафика является одной из актуальных задач при анализе городской среды ввиду их труднодоступности. Данные, получаемые из трафика, используются не только в задачах, связанных с

транспортным моделированием, но и в других смежных задачах – например, оценок массовости мероприятий или в оценках инвестиционной привлекательности городской среды. Традиционными методами подсчета и анализа городского трафика являются методы, в основном связанные либо с применением специализированного оборудования - например автоматизированных счетчиков, которые устанавливаются вдоль дорог или закладываются под асфальт, либо это натурные исследования и подсчеты. Типичными недостатками этих методов являются их дороговизна, труднодоступность и малоэффективность. Использование камер городского видеонаблюдения может выступать заменой натурным подсчетам, однако без вовлечения человека, круглосуточным и эффективным методом подсчета и анализа городского трафика.

Данный метод основан на отслеживании всех объектов на протяжении всего времени их движения в кадре. В процессе разработки было выявлено, что из-за качества камер и изменчивых погодных условий объекты периодически «пропадали». Для минимизации потерь данных было решено вести подсчет по линиям, отражающим характерные участки трафика, например, тротуар и автомобильные дороги, это, в свою очередь, позволит расширить подход и анализировать конкретные участки изображений – например, двери, разделять автомобильный трафик по полосам движения, также это позволяет полностью исключить повторную регистрацию одних и тех же объектов.

Для распознавания объектов было решено использовать свёрточную нейронную сеть YOLO 3 версии, так как это на данный момент одна из самых передовых нейросетей по скорости и точности [1]. YOLO была заранее предобучена на открытом наборе данных COCO, который содержит 80 типов объектов, среди которых нужны для задачи люди, автомобили, велосипеды, мотоциклы, автобусы и грузовики.

Для отслеживания движения объектов была взята идея алгоритма SORT, но без использования фильтра Калмана для прогнозирования движения [2]. Это позволит упростить алгоритм, а также сосредоточиться на регистрации реальных городских данных. Задача состоит в том, чтобы описать объектную модель, то есть представление объекта, используемое для передачи идентичности в следующий кадр. Состояние каждого объекта моделируется следующим вектором (1):

$$X = [u, v, s, r], \quad (1)$$

где:  $u$  и  $v$  – координаты центроида объекта,  $s$  – размер,  $r$  – соотношение сторон ограничивающей рамки. Для выполнения ассоциаций SORT использует математические эвристики, такие как максимизация показателей IoU (пересечение по объединению) между ограничивающими рамками в соседних кадрах. Каждое поле помечено номером (идентификатор объекта), и, если в следующем кадре нет соответствующего поля, алгоритм предполагает, что объект покинул кадр [3]. На рисунке показан пример работы алгоритма на реальных уличных камерах.

В ходе работы рассмотрена возможность применения камер городского видеонаблюдения в качестве источника данных о городском трафике, разработан подход и программная реализация, позволяющая получать данные о городском трафике в геоформате.

В дальнейшей работе планируется исследование непосредственных данных трафика и сопоставления их с косвенными источниками такими как - записи в социальных сетях, комментарии и отзывы в картографических сервисах (рисунок).

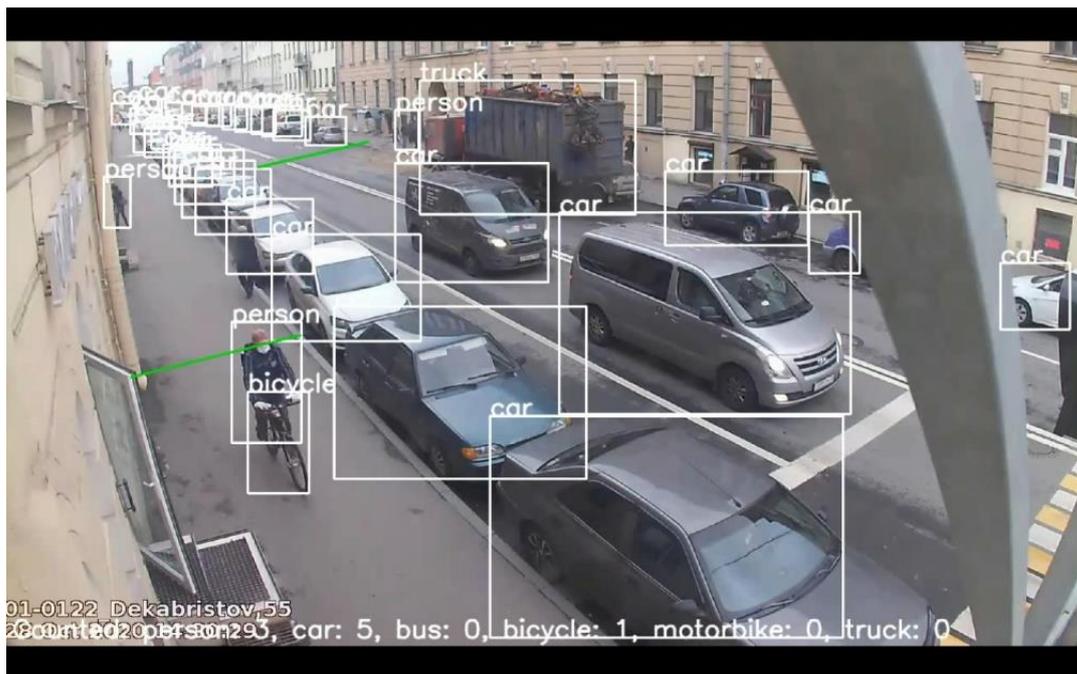


Рисунок. Демонстрация работы алгоритма

### Литература

1. Redmon J., Divvala S., Girshick R., Farhadi A. You only look once: Unified, real-time object detection // arXiv preprint arXiv:1506.02640, 2015.
2. Bewley A., Ge Z., Ott L., Ramos F. and Upcroft B., "Simple online and realtime tracking," 2016 IEEE International Conference on Image Processing (ICIP), Phoenix, AZ, 2016, pp. 3464-3468, doi: 10.1109/ICIP.2016.7533003.
3. Tracking Cows with Mask R-CNN and SORT | by Neuromation | Medium [Электронный ресурс]: электрон. статья, 2018. Режим доступа: <https://medium.com/neuromation-blog/tracking-cows-with-mask-r-cnn-and-sort-fcd4ad68ec4f> (дата обращения 09.01.2021).

**Чурякова Татьяна Алексеевна**

Университет ИТМО,  
Институт дизайна и урбанистики,  
студент группы № С42011,  
направление подготовки: 27.04.07 – Наукоемкие технологии  
и экономика инноваций,  
e-mail: tanufry@gmail.com

**Якимук Илья Витальевич**

Университет ИТМО,  
Институт дизайна и урбанистики,  
аспирант группы № 7738,  
направление подготовки: 05.13.10 – Управление в социальных  
и экономических системах,  
e-mail: ilyayakimuk@gmail.com

**Митягин Сергей Александрович**

Университет ИТМО,  
Мегафакультет трансляционных информационных технологий,  
Институт дизайна и урбанистики,  
к.т.н.,  
e-mail: mityagin@itmo.ru

УДК 004.04

**МЕТОД ОЦЕНКИ ПОТЕНЦИАЛА ЗЕЛЕННЫХ  
ЗОН ДЛЯ СОЗДАНИЯ ДИСПЕРСНЫХ ПАРКОВ**

**Т.А. Чурякова, И.В. Якимук**

**Научный руководитель – к.т.н. С.А. Митягин**

Работа выполнена в рамках темы НИР №620181 «Средства определения тематики и значимости объектов городской среды на основе открытых данных социальных сетей».

**Аннотация**

В статье рассматривается метод оценки потенциала озелененной территории для создания дисперсного парка. Для оценки потенциала территорий, подходящих для создания дисперсного парка, было рассмотрено две группы критериев: градостроительные характеристики озелененных территорий в условиях сложившейся застройки, наличие и характер объектов рекреационной и культурно-досуговой деятельности. По результатам исследования выявлен процент готовности муниципальных образований района для создания на их территории дисперсного парка. Итоговая оценка потенциала представлена по каждому из критериев, что позволяет оценить сильные и слабые места территорий.

**Ключевые слова**

Парк, дисперсный парк, общественные пространства, озелененные территории, управление парками, благоустройство, оценка потенциала, метод оценки.

Дисперсный парк - разновидность кластера, многофункциональное общественное пространство, сформированное из небольших озеленённых общественных пространств и объектов рекреационной и культурно-досуговой деятельности, объединённое линейными пешеходными пространствами [1].

Целью работы является оценка потенциала озелененных территорий для создания дисперсного парка в существующих градостроительных условиях на примере Калининского района города Санкт-Петербург.

Объектом исследования являются муниципальные округа Калининского района города Санкт-Петербург и ЗНОПы городского и местного значения.

Максимальное количество баллов для определения готовности создания дисперсного парка на территории – 100 баллов. Распределение баллов по критериям и подкритериям представлено в табл. 1.

Таблица 1

**Распределение баллов**

| Критерий                      | Результат                                       | Оценка, балл |
|-------------------------------|---|--------------|
| Градостроительный контекст    |   | 10           |
| Доступность                   | % горожан, не имеющих доступ                    | 5            |
| Функциональность              | % неохваченных функций                          | 5            |
| «Якорный» объект              |   | 20           |
| Площадь                       | % от максимально площади местного парка (10 га) | 3            |
| Функциональность              | % действующих функций                           | 8            |
| Проект благоустройства        | Да/нет  | 5            |
| Сообщества в социальных сетях | Да/нет  | 4            |
| Связность                     |   | 25           |
| Доступность                   | % горожан, имеющих доступ                       | 25           |
| Объекты культуры              |   | 15           |
| Функциональность              | % дополнительных функций                        | 15           |
| Объекты спорта                |   | 15           |
| Функциональность              | % дополнительных функций                        | 15           |
| Уникальные объекты            |   | 15           |
| Функциональность              | % дополнительных функций                        | 15           |

Потенциал территории критерия градостроительного контекста определяется исходя из процента числа жителей данного муниципального округа, не имеющих доступ к городским и районным паркам с развитой инфраструктурой [2-4]. Зона доступности представлена на рис. 1.

При этом в расчете потенциала учитывается количество функций, которые не представлены в доступных парках, как возможность реализации этих функций при создании дисперсного парка. Оценка критерия представлена в табл. 2.

При выборе «якорного» объекта потенциального дисперсного парка учитывается текущий градостроительный контекст: площадь территории, процент функций, представленных на территории. Для понимания вместимости парка и числа социальных групп, чьи потребности уже обеспечиваются в парке. Значимость территории для горожан рассчитывается исходя из наличия проектов благоустройства на территорию или проектов, реализованных за последние 5 лет, наличия сообществ в социальных сетях. Контекст территории представлен на рис. 2.

Оценка критерия представлена в табл. 3.

При расчете критерия связности территорий учитывается количество озелененных территорий, находящихся в пределах 5-минутной пешеходной доступности друг от друга [3]. Потенциал данного союза рассчитывается исходя из процента горожан, проживающих в этих зонах доступности [4]. Оценка критерия представлена в табл. 4.

Объекты рекреационной и культурно-досуговой деятельности представлены в трех сферах: культуры, спорта и уникальные объекты данной территории [1].

Учитываются объекты, находящиеся в уже определенной на прошлом этапе зоне доступности связанных озелененных территорий. Потенциал включения такого объекта в дисперсный парк определяется исходя из процента новых функций, которые могут разнообразить существующий досуг. Оценка критерия представлена в табл. 5.

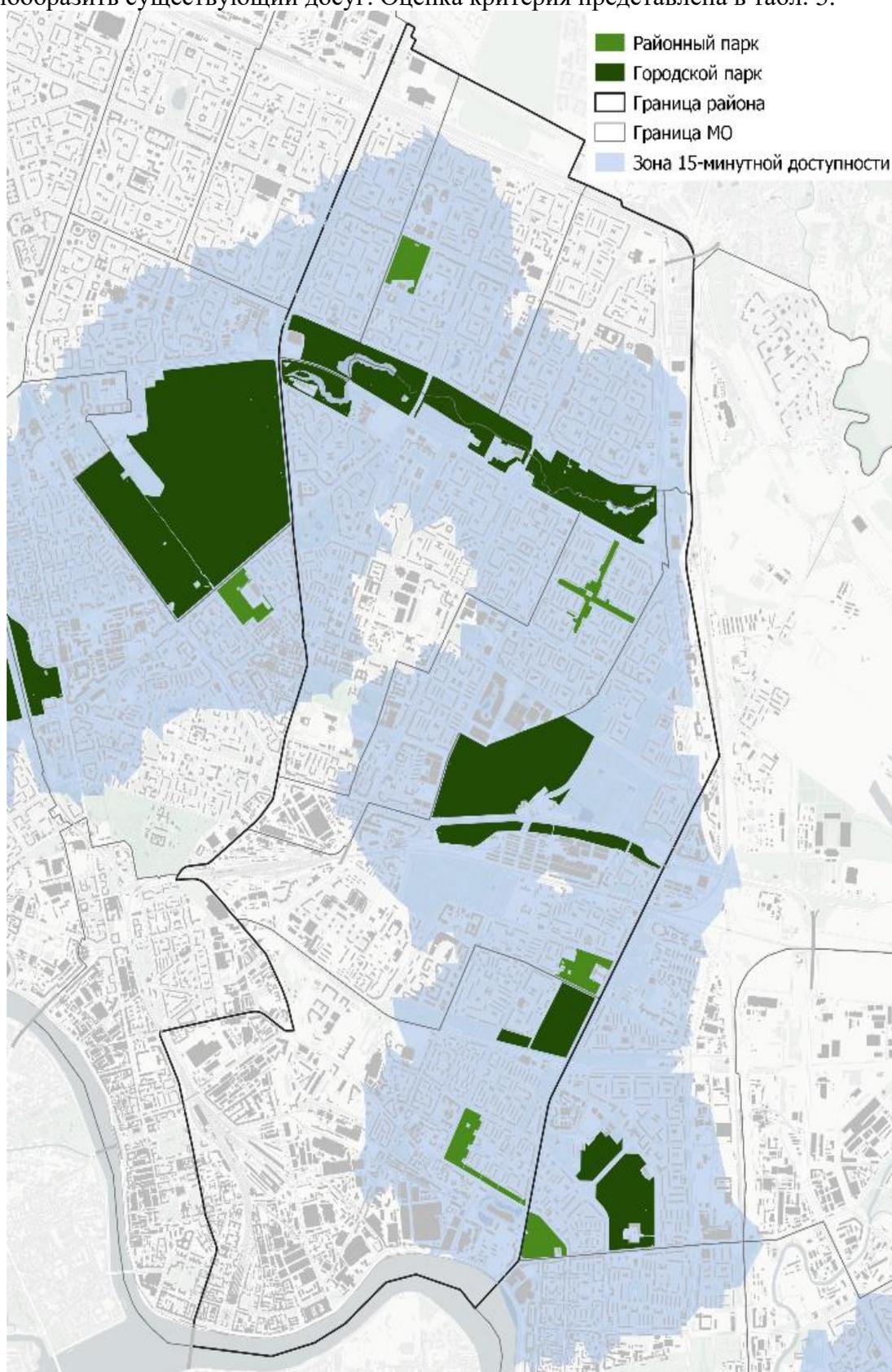


Рис. 1. Зона доступности городских и районных парков

**Градостроительный контекст**

| МО                   | Численность населения | Горожане, не проживающие в зоне доступности |    | Неучтенные функции |    | Баллы |
|----------------------|-----------------------|---|----|--------------------|----|-------|
|                      |                       | Чел.  | %  | Шт.                | %  |       |
| МО №21               | 79525                 | 35733                                       | 45 | 4                  | 24 | 3,4   |
| МО Академическое     | 108361                | 30317                                       | 28 | 2                  | 12 | 2,0   |
| МО Гражданка         | 74807                 | 1002  | 1  | 4                  | 24 | 1,3   |
| МО Пискаревка        | 62695                 | 34213                                       | 55 | 6                  | 35 | 4,5   |
| МО Прометей          | 74361                 | 17438                                       | 23 | 3                  | 18 | 2,1   |
| МО Северный          | 54106                 | 3172  | 6  | 3                  | 18 | 1,2   |
| МО Финляндский округ | 75332                 | 26268                                       | 35 | 6                  | 35 | 3,5   |

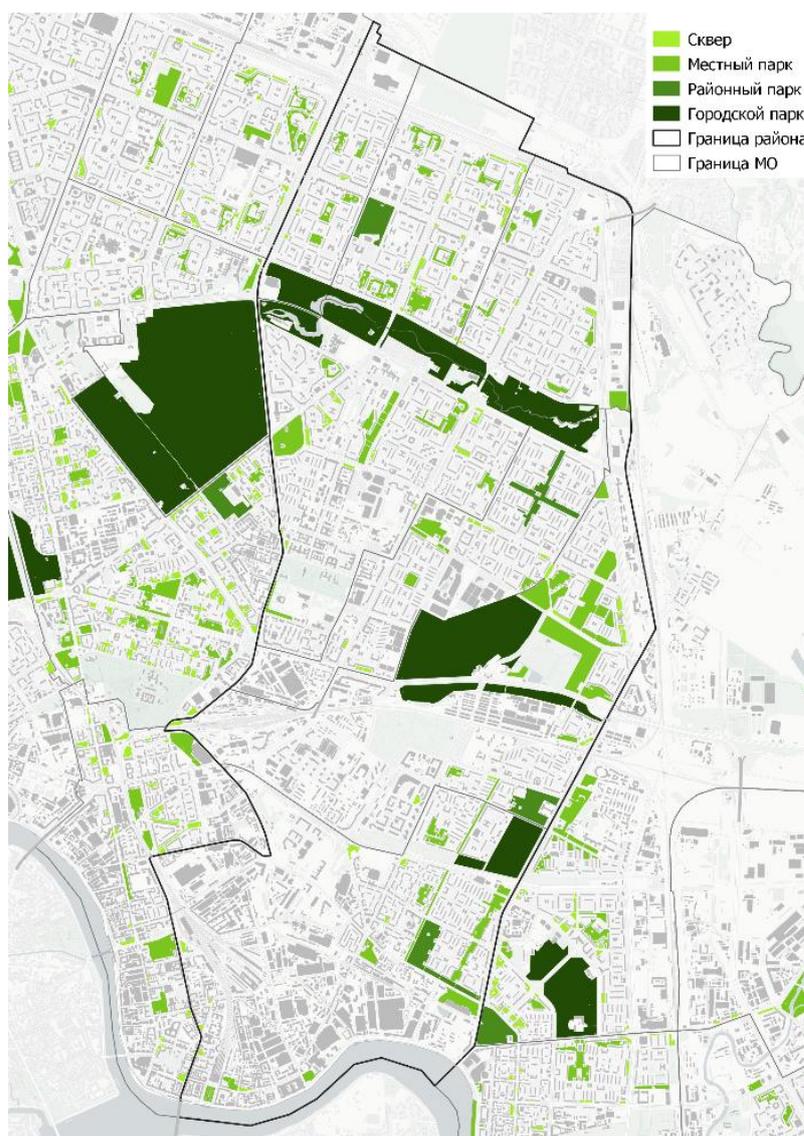


Рис. 2. Контекст

Таблица 3

**«Якорный» объект**

| МО                   |   | Площадь, га | Новые функции, шт | Проект благоустройства | Сообщества | Баллы |
|----------------------|---|-------------|-------------------|------------------------|------------|-------|
| МО №21               | Местный парк [2]<br>сквер на ул. Руставели восточнее пересечения с пр. Луначерского | 3,63        | 2                 | Нет                    | Нет        | 2,0   |
| МО Академическое     | сад Бенуа   | 9,85        | 9                 | Нет                    | Да         | 12,2  |
| МО Гражданка         | сад б/н на пересечении Гражданского пр. и ул. верности                              | 4,17        | 4                 | Нет                    | Нет        | 3,1   |
| МО Пискаревка        | сад б/н на пересечении ул. Руставели и пр. Науки                                    | 4,58        | 3                 | Да                     | Нет        | 5,4   |
| МО Прометей          | сквер б/н на Светлановском пр. д. 93  | 3,92        | 5                 | Да                     | Нет        | 8,5   |
| МО Северный          | Отсутствует   | -           | 0                 | Нет                    | Нет        | 0,0   |
| МО Финляндский округ | Сквер Безбородко  | 4,85        | 6                 | Нет                    | Нет        | 4,3   |

Таблица 4

**Связность территорий**

| МО                   | Численность населения, чел. | Количество территорий, связанных между собой зонами доступности | Горожане, проживающие в зоне доступности |     | Баллы |
|----------------------|-----------------------------|---|--|-----|-------|
|                      |                             |   | чел.                                     | %   |       |
| МО №21               | 80335                       | 0   | 1757                                     | 2   | 0,5   |
| МО Академическое     | 109393                      | 25  | 38309                                    | 35  | 8,8   |
| МО Гражданка         | 75558                       | 14  | 25020                                    | 33  | 8,3   |
| МО Пискаревка        | 62825                       | 24  | 17056                                    | 27  | 6,8   |
| МО Прометей          | 74988                       | 45  | 69201                                    | 92  | 23,1  |
| МО Северный          | 54569                       | 28  | 54574                                    | 100 | 25,0  |
| МО Финляндский округ | 75929                       | 27  | 34226                                    | 45  | 11,3  |

Таблица 5

**Объекты рекреационной и культурно-досуговой деятельности**

| МО                   | Объекты культуры |      | Объекты спорта |      | Уникальные объекты |      |
|----------------------|------------------|------|----------------|------|--------------------|------|
|                      | %                | балл | %              | балл | %                  | балл |
| МО №21               | 0                | 0    | 40             | 6    | 0                  | 0    |
| МО Академическое     | 60               | 9    | 40             | 6    | 0                  | 0    |
| МО Гражданка         | 0                | 0    | 40             | 6    | 33                 | 5    |
| МО Пискаревка        | 60               | 9    | 40             | 6    | 0                  | 0    |
| МО Прометей          | 40               | 6    | 80             | 12   | 17                 | 3    |
| МО Северный          | 0                | 0    | 60             | 9    | 17                 | 3    |
| МО Финляндский округ | 80               | 12   | 40             | 6    | 0                  | 0    |

Итоговая оценка по всем критериям представлена в табл. 6.

Таблица 6

**Результаты эксперимента**

| МО                   | Контекст | "Якорный" объект | Связность | Объекты культуры | Объекты спорта | Уникальные объекты | ИТОГО |
|----------------------|----------|------------------|-----------|------------------|----------------|--------------------|-------|
| МО №21               | 3,4      | 2,0              | 0,5       | 0                | 6              | 0                  | 12,0  |
| МО Академическое     | 2,0      | 12,2             | 8,8       | 9                | 6              | 0                  | 37,9  |
| МО Гражданка         | 1,3      | 3,1              | 8,3       | 0                | 6              | 5                  | 23,7  |
| МО Пискаревка        | 4,5      | 5,4              | 6,8       | 9                | 6              | 0                  | 31,6  |
| МО Прометей          | 2,1      | 8,5              | 23,1      | 6                | 12             | 3                  | 54,2  |
| МО Северный          | 1,2      | 0,0              | 25,0      | 0                | 9              | 3                  | 37,7  |
| МО Финляндский округ | 3,5      | 4,3              | 11,3      | 12               | 6              | 0                  | 37,0  |

В ходе данной работы была апробирована методика оценки потенциала зеленых зон для создания дисперсных парков. Исследуемая территория – семь муниципальных округов Калининского района. Оценка каждого округа проводилась на основании открытых данных (РГИС, OSM, Росстат) и натурного обследования. Для оценки доступности использовался сервер Galton, результат обрабатывался в программе QGIS Desktop. Результаты методики показали слабую готовность для округов Калининского района для создания дисперсных парков. Это может быть связано с наличием большого количества городских и районных парков в районе и на соседних территориях и их зоной доступности до горожан.

В ходе дальнейшего исследования планируется уточнить критерии оценки и опробовать метод на всей территории города Санкт-Петербург.

**Литература**

1. Ложкин А., Смирнова М., Разживина Д., Голодяев К., Гижицкая С., Карпов А., Скурихина Т., Архипова А., Юрченко Л., Лебедев Д., Дубынин А. Зеленый Новосибирск. Новосибирск: Вояж. 2017.
2. ДОМ РФ, STRELKA КБ. Книга 4. Стандарт освоения свободных территорий. 2019.

3. Urbica/galton [Электронный ресурс] // GitHub: [сайт]. [2019]. Режим доступа: <https://github.com/urbica/galton/> (дата обращения: 06.01.2021).
4. Росстат. Численность населения Российской Федерации по муниципальным образованиям // Федеральная служба государственной статистики. 2020. Режим доступа: <https://rosstat.gov.ru/compendium/document/13282> (дата обращения: 5.01.2021).



**Щепина Евгения Константиновна**

Год рождения: 1998  
Университет ИТМО,  
факультет цифровых трансформаций,  
студент группы № J41332с,  
направление подготовки: 01.04.02 – Прикладная  
математика и информатика,  
e-mail: jannyss@mail.ru



**Егорова Евгения Владимировна**

Год рождения: 1998  
Университет ИТМО,  
факультет цифровых трансформаций,  
студент группы № С42101,  
направление подготовки: 11.04.02 - Инфокоммуникационные  
технологии и системы связи,  
e-mail: evegorova@itmo.ru



**Федотов Павел Андреевич**

Год рождения: 1999  
Университет ИТМО,  
факультет цифровых трансформаций,  
студент группы № J41332с,  
направление подготовки: 01.04.02 – Прикладная  
математика и информатика,  
e-mail: pavelfedotov99@gmail.com

УДК 004.8

**АНАЛИЗ ИНТЕРЕСОВ И КУЛЬТУРНЫХ  
ПРЕДПОЧТЕНИЙ ПОЛЬЗОВАТЕЛЕЙ СОЦИАЛЬНЫХ  
СИСТЕМ НА ПРИМЕРЕ ИХ МУЗЫКАЛЬНЫХ ПОДПИСОК  
И ИХ ВЗАИМОСВЯЗЬ С ПСИХОТИПОМ**

**Е.К. Щепина, Е.В. Егорова, П.А. Федотов  
Научный руководитель – к.т.н., н.с. А.Г. Суриков**

Работа выполнена в рамках НИР №620174 «Инкрементальное обучение моделей психоэмоциональной динамики пользователей и сообществ в цифровых социальных системах».

**Аннотация**

Пользователи оставляют большое количество цифровых следов в цифровых социальных системах, к примеру, каждую минуту отправляется 41 миллион сообщений [4]. Значительная часть цифровых следов может быть представлена в топологии взаимосвязи: подписки на группы, музыка и пр. В данной работе мы предлагаем метод для анализа музыкальных интересов пользователей через построение графа музыкальных интересов и применение на нем методов неконтролируемого машинного обучения.

Целью данной работы является получение новых предикторов, которые можно использовать в качестве эмбедингов для решения задачи психоэмоционального типирования пользователей в социальных сетях. Результаты данного исследования могут иметь также самостоятельную ценность в исследовании социальных явлений в цифровом пространстве.

#### **Ключевые слова**

Анализ социальных сетей, теория графов, обработка данных, кластеризация графов, большая пятерка.

#### **Введение**

Задача психоэмоционального типирования пользователей в цифровом пространстве объединяет в себе две науки: психологию и науку о данных. Данная задача широко изучается в научном сообществе, и стандартным подходом решения является анализ взаимосвязи личностных характеристик пользователей и их данных, указанных в профилях в социальных сетях.

Мы решили проверить гипотезу, что культурные предпочтения пользователя значимо коррелируют с его психоэмоциональными особенностями.

Этапы работы по проверке гипотезы: было проведено анкетирование пользователей по модели «Большой пятерки» [1]; собраны данные по музыкальным интересам пользователей, прошедших анкетирования; проанализированы собранные данные и построен граф музыкальных интересов; кластеризован граф несколькими методами и произведен анализ полученных кластеров.

#### **Литературный обзор**

Использование социальных сетей и цифровых следов их пользователей для получения дополнительной информации о них - относительно новое научное направление, результаты исследований в котором могут быть применимы для решения множества реальных задач. Так, лучшее понимание личности человека позволяет эффективнее таргетировать для него рекламу и персонализировать настройки умных вещей, регулируя их в соответствии с характером человека. Возможно также и автоматическое выявление отдельных групп пользователей, например, склонных к депрессии или алкоголизму, по их профилям в социальных сетях, и оказание им соответствующей помощи.

Задачи, стоящие перед исследователем в этой области, зачастую носят неформализованный характер. Учитывая также высокую степень разнообразия данных (от текстовой информации - постов и комментариев - до изображений и видео пользователей), становится понятно, что решение подобного рода задач может быть достигнуто с помощью множества методов, в корне отличающихся друг от друга: классификация, регрессия, нейронные сети.

В случаях, когда данные пользователей не размечены, используются различные варианты кластеризации информации, также имеющей множество форм применения. В работе [4] твиты пользователей были предобработаны, а затем построена матрица силы взаимосвязи пользователей в соответствии с использовавшимися ими словами. К полученным данным были применены обычная и спектральная кластеризации методом К-средних, суть последней из которых заключается в том, чтобы использовать собственные значения полученной матрицы силы взаимосвязи для осуществления снижения размерности перед кластеризацией в пространствах уже меньших размерностей. Результаты работы обоих видов кластеризации оказались практически идентичными, однако применять спектральную кластеризацию возможно лишь при наличии больших вычислительных мощностей из-за ее высокой вычислительной сложности.

Пользовательские данные нередко представляются различными способами в виде графов, работа с которыми способна повысить качество итоговых результатов. Подобный подход к работе с данными из социальных сетей нередко встречается в научной литературе. В публикации авторы представили данные о сообществах в социальной сети LiveJournal в виде графа, который в последующем был кластеризован с помощью агломеративной спектральной кластеризации. Результаты распознавания сообществ таким методом продемонстрировали значение F-меры, равной 75%, что оказалось больше значения F-меры другого алгоритма AN-KSC, также примененного к этим данным, на 14 процентных пунктов. В другой работе ученые представили пользователей и их отношения в социальной сети Sina Weibo в виде взвешенного неориентированного графа, вершины в котором представляли пользователей, а ребра - отношения между ними. Полученный граф был кластеризован с помощью метода агломеративной кластеризации ABCD и алгоритма Клаузета-Ньюмана-Мура. По результатам их сравнения был сделан вывод о том, что алгоритм ABCD справился с поставленной задачей лучше, так как смог определить менее крупное сообщество социальной сети ввиду принятия во внимание большего числа факторов сети.

Как мы можем заметить, работа с графами зачастую требует применения особых алгоритмов кластеризации, разработанных специально для графовых структур. Одними из вариантов таких алгоритмов, помимо вышеназванных, являются алгоритмы Affinity Propagation, Louvain и Leiden, предназначенных для выявления сильносвязных сообществ в социальной сети. Используемый в данной статье метод Affinity Propagation формирует кластеры благодаря посылке сигналов между вершинами графа. В результате данной операции выявляются более и менее показательные для графа вершины. Те, что характеризуют данные лучшим образом, остаются, другие же исключаются из рассмотрения. Данная процедура повторяется до тех пор, пока не будет достигнута точка сходимости, в которой известны кластеры для всех вершин графа.

### Данные

Для решения задачи психотипирования пользователей было проведено сравнительное исследование существующих методик тестирования. При участии эксперта-психолога была выбрана модель «Большой пятерки», как наиболее научно-обоснованная и часто используемая в научных исследованиях методика.

Для анкетирования пользователей было разработано приложение «VK Tests» [5] для VK Mini Apps. В результате проведённого анкетирования были получены данные 978 пользователей, среди которых только у 306 удалось извлечь их музыкальные предпочтения. Это пользователи обоих полов, проживающие преимущественно в России, а также небольшое количество жителей стран СНГ, возраст пользователей разнится от 18 до 67 лет, однако большинство опрошенных в возрасте 18-30 лет. Всего было получено 53556 музыкальных композиций.

### Построение модели

Модель интересов пользователей была сформирована в виде графа, вершины которого представляют собой музыкальных исполнителей, которые находятся в коллекциях у пользователей социальной сети, а связи между вершинами отражают факт наличия у одного отдельно взятого пользователя композиции обоих музыкальных исполнителей. При этом веса ребер обозначают количество таких связей и могут быть представлены следующей формулой:

$$w(i, j) = \sum_{u \in U} V_i \cap V_j,$$

где  $U$  – множество пользователей,  $V_i$  – наличие композиций  $i$ -того музыкального исполнителя у пользователя  $u$ .

В соответствии с полученной формулой для веса ребра при увеличении этого веса вершины расположены семантически ближе друг к другу. В связи с этим, для дальнейшего корректного использования алгоритмов кластеризации веса ребер были заменены на обратные значения:

$$w^*(i, j) = \frac{1}{w(i, j)}.$$

Построенный неориентированный граф обладает следующими характеристиками (табл. 1).

Таблица 1

#### Характеристики построенного графа

| Характеристика | Количество вершин | Количество ребер | Диаметр сети | Плотность графа | Количество компоненто в связности | Средний коэффициент кластеризации |
|----------------|-------------------|------------------|--------------|-----------------|-----------------------------------|-----------------------------------|
| Величина       | 59                | 2846             | 3            | 0.832           | 1                                 | 0.926                             |

#### Результаты

В качестве алгоритмов кластеризации были использованы Affinity Propagation и Spectral Clustering. Для расчета близости было использовано евклидово расстояние между вершинами.

В результате кластеризации методом Affinity Propagation было выделено 13 кластеров; при использовании метода Spectral Clustering - 8 кластеров, также была рассмотрена кластеризация Spectral Clustering с заданным количеством кластеров равным 5.

Результаты кластеризации были оценены следующими метриками качества кластеризации: Modularity, Performance, Coverage, Silhouette index, Calinski-Harabasz index, Davies-Bouldin index (табл. 2).

Таблица 2

#### Оценка качества кластеризации

|                             | Modularity | Performance | Coverage | Silhouette index | Calinski-Harabasz index | Davies-Bouldin index |
|-----------------------------|------------|-------------|----------|------------------|-------------------------|----------------------|
| Spectral Clustering (n=5)   | -0.109     | 0.229       | 0.221    | 0.066            | 5.542                   | 2.267                |
| Spectral Clustering (n=8)   | -0.087     | 0.243       | 0.115    | 0.036            | 4.071                   | 2.238                |
| Affinity Propagation (n=13) | -0.063     | 0.311       | 0.197    | 0.058            | 4.646                   | 1.136                |

В соответствии с полученными метриками качества кластеризации, для дальнейшего анализа был выбран метод распространения близости (Affinity Propagation). Для кластеризации данным методом было визуализировано разбиение на кластеры на наборе данных (рис. 1).

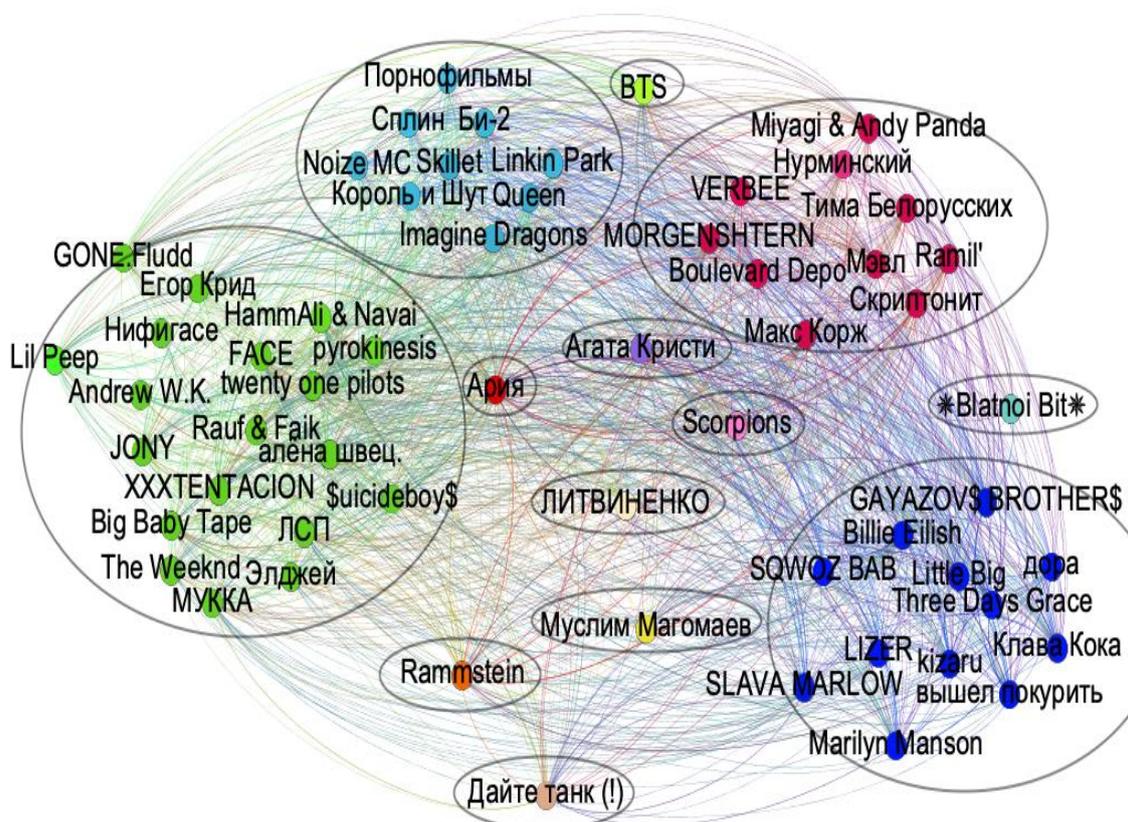


Рис. 1. Визуализация результатов кластеризации методом Affinity Propagation

Можно заметить, что некоторые исполнители были определены в отдельные кластеры. Эксперты интерпретировали такое разбиение в связи с сильным различием характера музыкальных композиций этих исполнителей со всеми остальными исполнителями.

Для определения соответствия полученных кластеров с диспозициями личности человека были построены линейные графики распределения величин пяти диспозиций модели человека. Величины каждой диспозиции были найдены следующим образом: для каждого кластера были определены исполнители, вошедшие в него, и для каждого такого исполнителя были определены средние величины пяти диспозиций: экстраверсии, доброжелательности, сознательности, нейротизма и открытости опыту по всем пользователям выборки, у которых в их музыкальной коллекции есть данный исполнитель. Результат анализа представлен на рис. 2.

На рис. 2 одинаковым цветом выделены музыкальные исполнители, относящиеся к одному кластеру. Можно заметить, насколько отличаются траектории исполнителей-кластеров: “BTS” и “Муслим Магомаев” друг от друга и других исполнителей. Исполнители были выбраны случайным образом, при этом было осуществлено несколько подобных случайных выборок, результаты среди которых схожи.

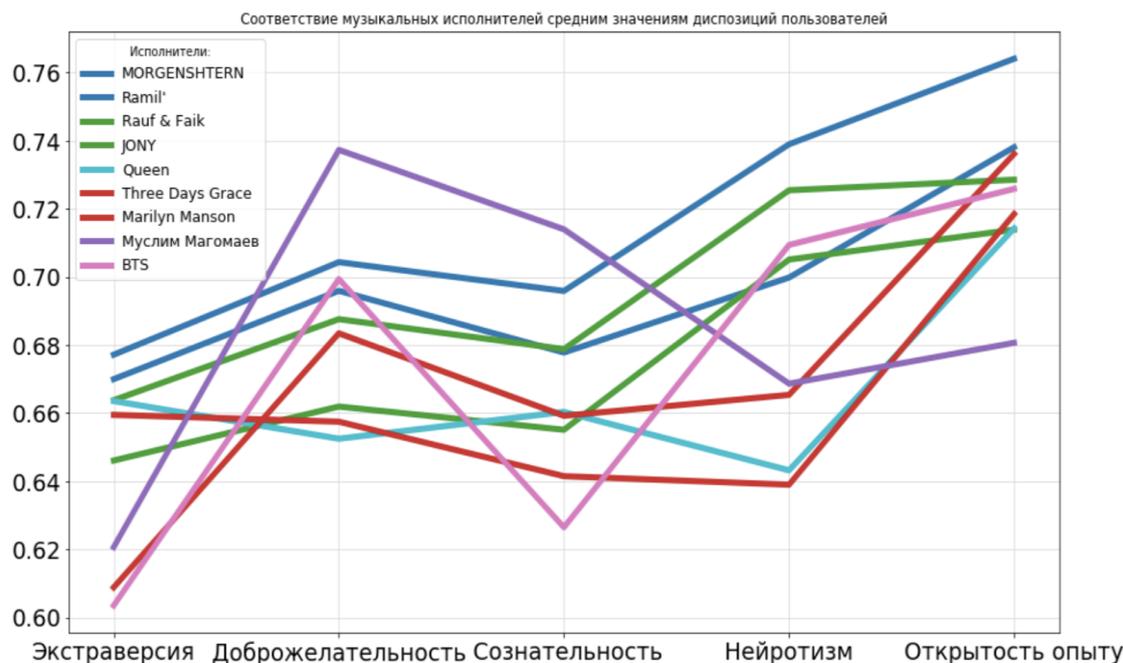


Рис. 2. Соответствие музыкальных исполнителей средним значениям диспозиций пользователей

Экспертно интерпретировав результаты, было определено, что пользователи, имеющие в своих музыкальных коллекциях исполнителей из одного кластера, имеют схожую траекторию диспозиций личности человека, в то время как у исполнителей из разных кластеров они различаются. Это может быть связано с различиями в психотипах пользователей, предпочитающих музыку различных кластеров. Например, пользователи, предпочитающие музыку исполнителей первого кластера (“MORGENSHTERN”, “Ramil”), имеют доминантную черту «Открытость опыту», что характеризуется непокорностью человека, что свойственно более молодым пользователям, у которых имеются такие композиции.

### Заключение и дальнейшая работа

Цель этого исследования состояла в том, чтобы получить новые предикторы, которые можно использовать в качестве эмбедингов для решения задачи психоэмоционального типирования пользователей в социальных сетях. Для достижения цели были собраны данные по музыкальным интересам пользователей, прошедших тестирование по модели «Большой пятерки» [1], построен граф музыкальных интересов пользователей в соответствии с изобретенной моделью, затем граф был кластеризован с помощью методов Spectral Clustering (Shi-Malik Algorithm) [2] и Affinity Propagation Clustering [3]. Наши результаты показали, что между культурными интересами пользователей и их психоэмоциональными особенностями существует взаимосвязь. В связи с повсеместным использованием социальных систем в жизни людей актуальность данной работы заключается в изучении феноменов поведения людей в этих системах как для понимания их природы, так и для дальнейшего улучшения рекомендательных алгоритмов.

В дальнейшем планируется перейти от унимодального пространства (музыкальные интересы) к мультимодальному пространству (музыкальные интересы, подписки на группы и т. д.). Для дальнейших исследований будет собрана более репрезентативная выборка данных пользователей и их интересов и культурных предпочтений, улучшена модель кластеризации графа и интерпретация его кластеров.

### Литература

1. Shi J., and Malik J., «Normalized cuts and image segmentation» // IEEE Transactions on Pattern Analysis and Machine Intelligence 22(8). 2000. P. 888–905.
2. Kuldeep S. & Harish Biswas S., Bhaskar. «Clustering of People in Social Network based on Textual Similarity» // Perspectives in Science. 8, 2016.
3. Lanel G.H.J. and Jayawardena H.S.S.P. «A Study on Graph Theory Properties of On-line Social Networks» // International Journal of Scientific and Research Publications. vol. 10. no. 3. 2020. P. 9929.
4. Ulzii-Utas Narantsatsralt, Sanggil Kang, «Social Network Community Detection Using Agglomerative Spectral Clustering» // Complexity. vol. 2017. 3719428. P. 1-10.
5. Ruifang L. & Shan F. & Ruisheng S. & Wenbin G. «Weighted Graph Clustering for Community Detection of Large Social Networks» // Procedia Computer Science. 31. 2014. P. 85–94.



**Эльдиб Полина Юрьевна**  
Университет ИТМО,  
Институт дизайна и урбанистики,  
студент группы № С42042,  
направление подготовки: 09.04.03 – Прикладная информатика,  
e-mail: xsuperpo@gmail.com

УДК 004. 9

## **ФОРМИРОВАНИЕ ФУНКЦИОНАЛЬНЫХ ТРЕБОВАНИЙ К ГЕОИНФОРМАЦИОННЫМ СИСТЕМАМ ОБЩЕСТВЕННОГО УЧАСТИЯ**

**П.Ю. Эльдиб**

**Научный руководитель – к.соц.н., доцент А.Е. Ненько**

Работа выполнена в рамках темы НИР «Формирование функциональных требований к геоинформационным системам общественного участия».

### **Аннотация**

В работе описан процесс формирования функциональных требований к геоинформационным системам общественного участия, которые являются инструментом для сбора и обработки субъективного мнения горожан, а также могут выступать в качестве платформы для обсуждения идей территориального развития и предложения собственных идей. Сформулированные функциональные требования складываются из требований к процессу общественного участия, геоинформационным системам, интернет-технологиям и пользовательских требований, выявленных в ходе интервью с заинтересованными сторонами.

### **Ключевые слова**

Геоинформационные системы, общественное участие, интернет-технологии, геоинформационные системы общественного участия, функциональные требования, заинтересованные стороны, соучаствующее картирование.

С 2018 года начался федеральный проект «Формирование комфортной городской среды» [1]. Руководством страны был признан факт, что улучшение качества городской среды необходимо не только для эстетики и комфорта, но и для победы нашей страны в гонке за лучшими умами, идеями и инвестициями.

Общественное участие в территориальном планировании помогает выявить потребности и интересы заинтересованных сторон и достичь согласия по основным пунктам реализации проектов [2]. Вовлечение общественности является неотъемлемой частью надлежащего государственного управления. Общественное участие обеспечивает сотрудничество большего числа взглядов, потребностей и интересов, что ведет к более эффективному государственному управлению [3].

Интернет-инструменты и технологии предоставляют новые каналы взаимодействия с обществом, которые могут внести существенный вклад в проекты общественного участия.

Геоинформационные системы (ГИС) могут выступать как инструмент сбора общественно необходимой географической информации, инструмент обработки информации, и инструмент отображения интересов всех заинтересованных сторон.

Пересечение социальных процессов общественного участия, пространственных технологий ГИС и интернет-технологий называется сферой интересов географических информационных систем общественного участия. ГИСОУ (геоинформационная система общественного участия) – это метод и инструмент соучаствующего картирования, который направлен на вовлечение граждан в процесс принятия решений.

Взаимодействие и взаимосвязь между каждым двумя перекрывающимися сферами создают 3 поддомена: веб-общественное участие, веб-ГИС и ГИСОУ. Область, объединяющая все три сферы, под названием веб-ГИСОУ. На рис. 1 представлена схема, иллюстрирующая сферы интересов веб-ГИСОУ.



Рис. 1. Сферы интересов веб-ГИСОУ

Аналогично лестнице общественного участия Шерри Арнштейн [4] (рис. 2), исследователи из Манчестерского университета в Великобритании Кингстон и Карвер предложили лестницу электронного общественного участия [5] (рис. 2). Лестница электронного участия иллюстрирует различные формы онлайн-участия в зависимости от уровней участия. Эту модель можно применить к ГИСОУ, чтобы включить в нее различные уровни общественного участия, начиная от формальных мер (например, информирование граждан через веб-сайт), до партнерства (например, онлайн обсуждение и опросы общественного мнения, где ГИС может использоваться для выявления конфликтов), и даже до гражданского управления (например, онлайн-система поддержки принятия решений, где участники представляют свой выбор политики).

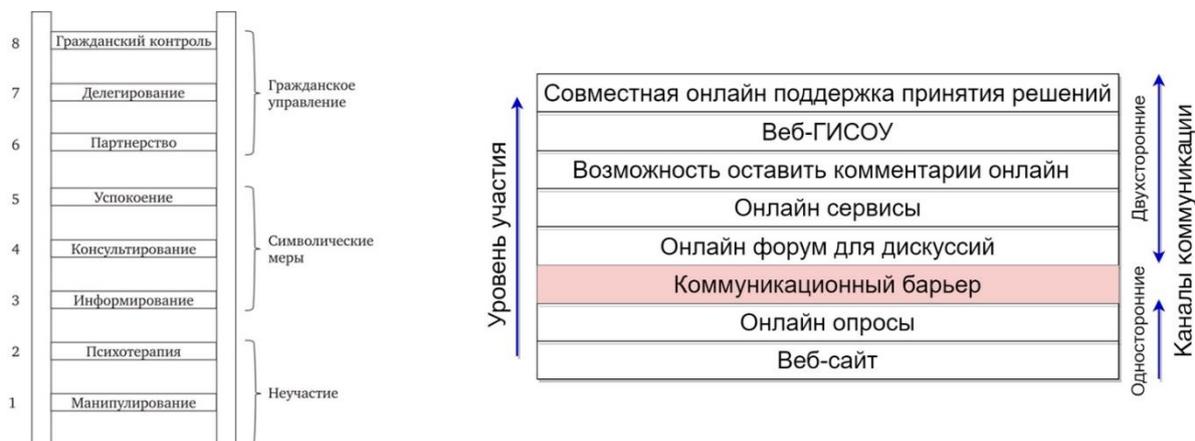


Рис. 2. Лестница гражданского участия по Шерри Р. Арнштейн [4] и электронное общественное участие по Кингстону и Карверу [5]

Прежде чем сформулировать сценарии использования геоинформационных систем общественного участия и определить функциональные требования, необходимо выявить ключевые заинтересованные стороны.

На рис. 3 показана луковая диаграмма, которая определяет, какая заинтересованная сторона напрямую взаимодействует с ГИСОУ. Кольца построены на основе способности заинтересованных сторон влиять на систему.

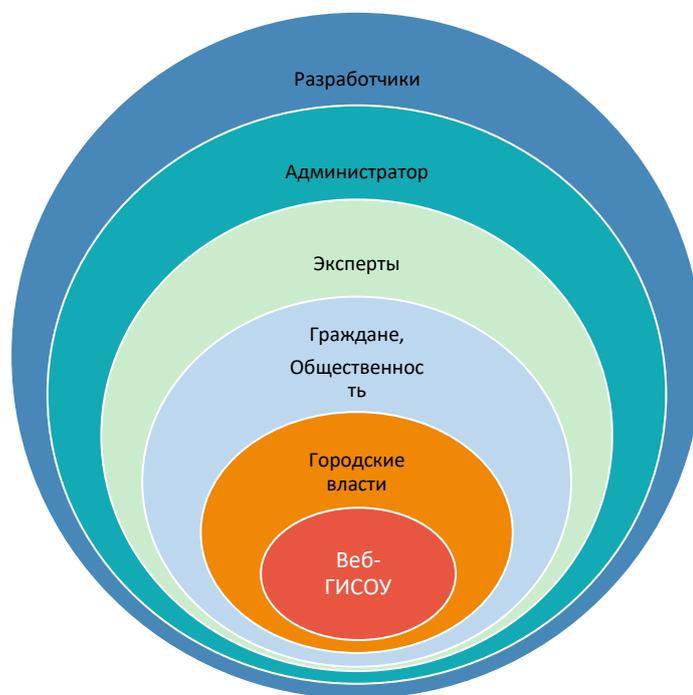


Рис. 3. Луковая диаграмма заинтересованных сторон

Органы городского планирования, которые отвечают за формулирование окончательной политики, имеют наиболее сильное влияние на систему. Далее идут граждане, которые, являются конечными пользователями системы. Затем идут участники, которые действуют как посредники и имеют некоторый интерес к системе. Затем администратор, который настраивает и контролирует веб-параметры и в конце команда разработчиков.

В рамках научно-исследовательской работы осенью-зимой 2020 года были проведены интервью с пользователями, разработчиками, экспертами по общественному участию. Интервью проводились по российским ГИСОУ Mapsurvey, Otmetky, LATL.NG. По итогу интервьюирования пользователей были составлены пользовательские требования к ГИСОУ и варианты использования. На рис. 4 показаны примеры диаграмм вариантов использования.

Контекстная диаграмма показывает, как данные обмениваются между внешними пользователями и системой. На рис. 5 показана контекстная диаграмма с основными функциями, участниками и идентификацией внешнего источника данных.

Принципы общественного участия, ГИС и заинтересованные стороны - это исходные данные на стадии разработки требований. Используя контекстную диаграмму, эти три составляющие образуют модель системы. Контекстная диаграмма связывает функции системы с внешним миром. Контекст веб-ГИСОУ будет представлять собой виртуальную платформу для совместной работы с доступом к ГИС в режиме онлайн. Платформа будет использоваться группами (оказывающими разное влияние на систему), участвующими в решении городских проблем (с обсуждениями, гео-опросами, предложениями инициатив, обратной связью и т. д.).

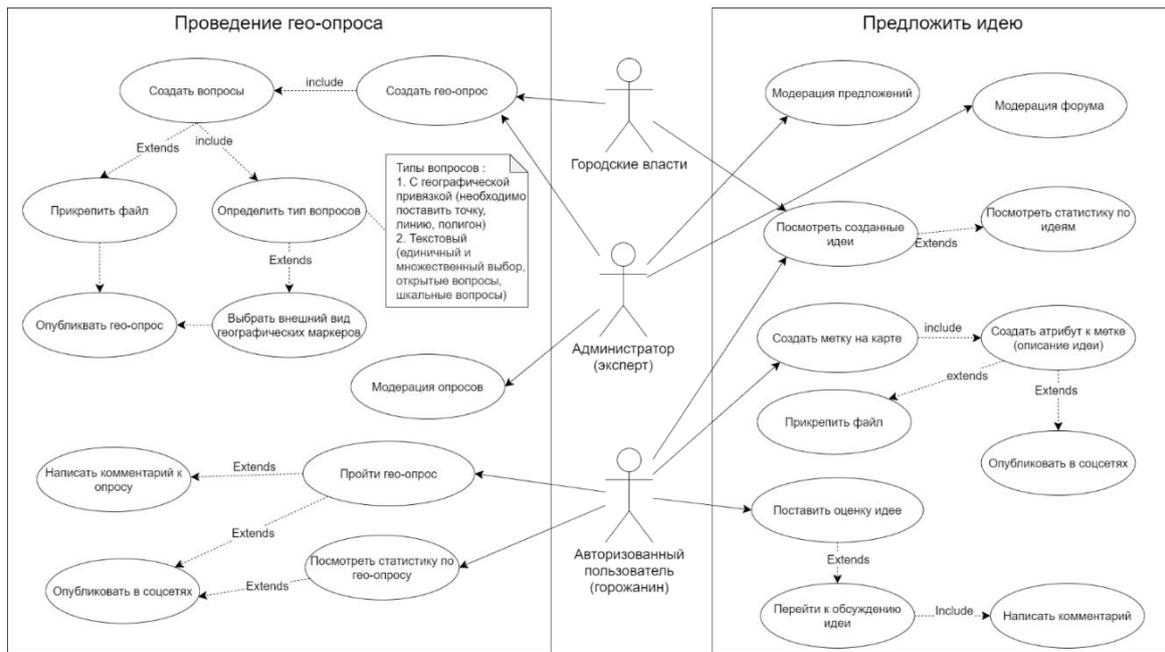


Рис. 4. Диаграммы вариантов использования



Рис. 5. Контекстная диаграмма веб-ГИСОУ

Традиционно любая географическая информационная система основана на принципах ввода, управления, анализа и представления данных. ГИС состоит из следующих компонентов: компонент для сбора географической информации, печатные

и цифровые карты, компонент базы данных, позволяющий хранение и манипулирование пространственными данными и визуализация, объединяющая слои с данными для пространственных аналитических операций. Система должна быть совместима со стандартами ISO, Open Geospatial Consortium и World Wide Web Consortium.

Типовая архитектура веб-ГИС системы в основном является «клиент-серверной» (рис. 6). Клиент — это приложение, которое может связываться с сервером через стандартный веб-протокол, например HTTP. Это приложение может быть в виде веб-браузера или отдельной утилиты. Веб-сервер отвечает за обработку запроса от клиента и доставку соответствующего ответа. В архитектуре веб-ГИС веб-сервер также взаимодействует с серверным компонентом ГИС.

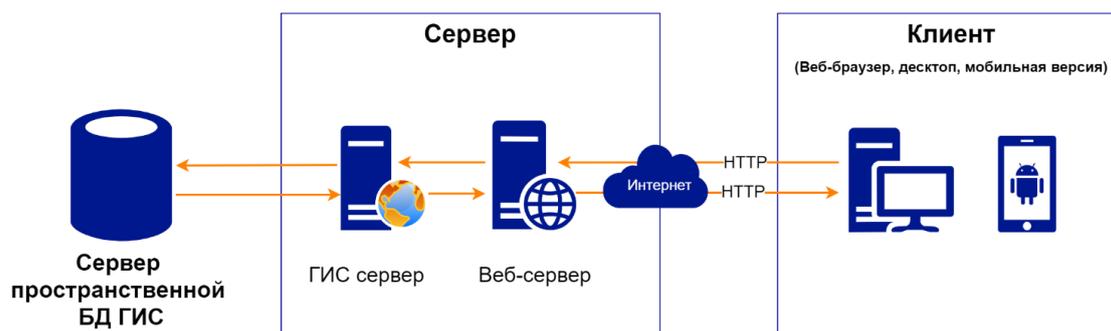


Рис. 6. Типовая архитектура веб-ГИС системы

Для дальнейшего проектирования ГИСОУ необходимо создать модель поведения проектируемой системы, в частности, алгоритмической детализации и логической реализации операций, которые может выполнять система. Для этого будет использоваться диаграмма деятельности UML.

Геоинформационные системы общественного участия – это молодая перспективная сфера. Одновременно в России появилось сразу несколько ГИСОУ (Marsurvey, Otmetky, LATL.NG), и сейчас они активно развиваются. Появилась конкурирующая среда, что говорит о том, что существует потребность в данном инструменте у горожан, активистов, экспертов и городских властей.

Данное исследование может послужить в качестве методических рекомендаций по описанию функциональных требований к геоинформационным системам общественного участия.

## Литература

1. Паспорт приоритетного проекта «Формирование комфортной городской среды» в действующей редакции.
2. Деменев А.Г. и др. Опыт общественного участия в планировании комфортной городской среды на примере Архангельской области // Арктика И Север. 2018. № 33.
3. Тайлер Д. Ноултон. Вовлечение общественности: построение институционального потенциала. Записки ИГУП. Выпуск №1, 2013 [Электронный ресурс]. Режим доступа: [https://www.ucentralasia.org/Content/Downloads/UCA-IPPA\\_OP1\\_Public\\_Engagement\\_Rus.pdf](https://www.ucentralasia.org/Content/Downloads/UCA-IPPA_OP1_Public_Engagement_Rus.pdf) (Дата обращения 18.02.2021).
4. Arnstein S.R. A Ladder Of Citizen Participation // Journal of the American Institute of Planners. 1969. 35:4. С.217.
5. Kingston R. The role of envirovernment and public participation in the planning process. В кн.: XVI AESOP Congress, Volos, Greece, 2002.



**Ян Ци**

Год рождения: 1992  
Университет ИТМО,  
факультет информационных технологий и программирования,  
аспирант,  
направление подготовки: 09.06.01 – Информатика  
и вычислительная техника,  
e-mail: yangqi@itmo.ru



**Фарсеев Александр Игоревич**

Университет ИТМО,  
факультет информационных технологий и программирования,  
PhD, профессор-исследователь,  
e-mail: farseev@itmo.ru

УДК 004.852

**TOWARDS MULTI-VIEW MULTI-SOURCE LEARNING FROM  
SOCIAL MULTIMEDIA FOR HUMAN PERSONALITY PROFILING**

**Ян Ци**

**Научный руководитель – PhD, профессор-исследователь А.И. Фарсеев**

**Аннотация**

Человеческие черты личности являются ключевыми факторами, определяющими принятие нами решений, ежедневно влияющими на наш жизненный путь. Выводы о таких чертах личности, как тип личности Myers-Briggs, а также понимание зависимостей между личностными качествами и поведением пользователей в различных социальных сетях имеют решающее значение для современных исследований и промышленных приложений. Появление разнообразных и многоцелевых социальных сетей позволяет автоматически и эффективно выполнять профилирование личности пользователя на основе много модальных данных из множества источников данных. Тем не менее, исследовательские усилия по профилированию личности на основе данных из множественных социальных сетей относительно скудны, а уровень влияния различных данных из социальных сетей на эффективность обучения на компьютере еще предстоит всесторонне оценить. Данное исследование является одной из первых попыток преодоления такого важного исследовательского пробела. В частности, в этой работе мы делаем вывод об индикаторе типа личности Myers-Briggs (Myers-Briggs Personality Type), применяя новую мульти-версию фреймворка под названием “MMPIS” и сравнивая результаты работы не только с различными способами передачи данных, но и с различными источниками данных в социальных сетях. Наши экспериментальные результаты демонстрируют способность MMPIS извлекать уроки из многопрофильных данных для профилирования личности путем эффективного использования значительной разницы в данных, поступающих из различных источников социальных мультимедиа. Мы также обнаружили, что выбор подхода к машинному обучению имеет решающее значение при выборе источников данных социальных сетей, и что люди склонны

раскрывать многогранность своей личности в различных социальных сетях.

### **Ключевые слова**

Машинное обучение, анализ социальных сетей, профилирование пользователей.

В течение последнего десятилетия быстро появляется все больше и больше социальных медиа-платформ, которые играют жизненно важную роль в содействии человеческому взаимодействию во всем мире. С 2012 года среднее ежедневное время просмотра в социальных сетях увеличилось с 60 минут до 144 минут. Более того, оно увеличилось еще больше с момента начала вспышки болезни COVID-19, когда люди были заперты дома с единственным оставшимся вариантом взаимодействия с друзьями через социальные сети.

Для поддержания высокого уровня вовлеченности пользователей необходимо, чтобы социальные сети позиционировали и рекомендовали соответствующий контент в соответствии с интересами и поведением пользователей. Например, экстраверты более склонны использовать социальные сети в целом, так как они имеют тенденцию раскрывать себя как энтузиастов, интерактивных, и, следовательно, формировать больше социальных кругов вокруг себя. Однако, например, было обнаружено, что интроверты тратят значительно больше времени на оценку стоимости каждого онлайн-сервиса, которым они пользуются, прежде чем может произойти более глубокое вовлечение в работу сервиса. Однако, в ручном режиме, с такими большими и разнообразными данными, доступными в настоящее время в социальных сетях, трудно отличить таких пользователей за предоставление им более персонифицированного опыта. И поэтому необходим автоматизированный подход к пониманию моделей человеческого поведения в социальных сетях. К сожалению, в настоящее время профилирование личности в значительной степени зависит от ручных анкет, и поэтому его стоимость остается неприемлемо высокой, ограничивая его использование в персонифицированных сервисах реального времени, таких как социальные сети [3-5].

В связи с этим возникла необходимость знать личность в большом масштабе в автоматическом режиме, и о её сложности широко сообщалось, так как модели поведения пользователей и характеристики между платформами социальных медиа варьируются. Например, Twitter в основном используется для обмена идеями, в то время как Facebook используется для общения с друзьями и семьей. Кроме того, Facebook имеет больше аудиторий различных возрастных групп, в то время как аудитория Twitter в основном составляет от 18 до 29 лет. Кроме того, по сравнению с Facebook, Twitter концентрируется на выражении мнения пользователя, а не на его личности. Несмотря на то, что пользователи Twitter или Facebook могут проводить время на PersonalityCafe, личностно-ориентированном форуме, в котором темы гораздо больше концентрируются на личностных особенностях. Отсутствие социального давления, вызванного анонимностью и степенью концентрации тем, может выявить различные особенности использования среди Facebook и Twitter. Поэтому важно иметь представление о влиянии автоматического определения личностных характеристик на различных социальных медиа-платформах.

За последние два десятилетия ряд исследователей стремились профилировать личностные характеристики пользователей с помощью автоматического подхода. Возникшая из статистического анализа английской лексики, модель Big Five раскрывает тесную взаимосвязь между личностью и языком, заложила основу автоматического профилирования личности, поскольку она объединяет прогнозирование личности и использование языка. Исследования на ранней стадии автоматического профилирования личности на основе небольших наборов данных [11]. После чрезвычайной ситуации проекта MyPersonality [9], первого крупномасштабного набора данных, использующего данные, генерируемые

пользователями из платформы социальных сетей с метками личности пользователей в Facebook, автоматическое профилирование личности на основе данных из социальных сетей начинает привлекать внимание исследовательского сообщества [10, 7, 4]. Однако в большинстве предыдущих исследований использовался только один текстовый метод и извлекался лексикон, например, Linguistic Inquiry and Word Count (LIWC), предварительно подготовленные текстовые вставки, такие как Global Vector for Word Representation (GloVe) [3], или словарный подход, основанный на частоте слов, например TF-IDF.

К сожалению, мало что было сделано для получения пользы от использования всех возможных способов получения данных в задании профилирования личности, несмотря на то, что в последние годы все больше внимания уделяется многопрофильному обучению [2]. Исследования доказали, что вместо того, чтобы использовать единую модальность, использование преимуществ множественных перспектив точки данных значительно увеличивает производительность модели. Кроме того, в связи с отсутствием исследований в области сравнения влияния выбора платформы социальных медиа, разница в поведении пользователей между платформами социальных медиа остается неясной. В нашем исследовании мы извлекаем семантические свойства как из текста, так и из изображения из своего аккаунта в социальных сетях, чтобы получить полное представление о профилировании личности.

Основной вклад данного исследования состоит из двух частей: во-первых, мы представляем много просмотренный фреймворк вывода личности и раскрываем различия в профилировании личности среди платформ социальных сетей; во-вторых, мы выпустили большой масштабный мульти-источник мульти-модальных данных с меткой личности, который станет новым эталоном для много источниковых мульти просмотренных задач профилирования личности.

### **Multi-View индивидуальность данных мульти-модальности**

Данные собирались с 1 января 2018 года по 1 января 2021 года. Этапы описаны ниже.

1. Сбор истины на земле. Для выполнения задачи по сбору истины в "Twitter" мы просматривали твиты в период с 1 января 2019 года по 1 января 2021 года, которые содержат такие ключевые слова, как "I'm an ENTP" или "I am an ENTP", чтобы получить список пользователей в качестве истины в Twitter, как правило, после теста на личность публикуются самоотчеты о личностных признаках из сообщений пользователей. Между тем, мы отмечаем, что пользователи оставляют положительный комментарий со своим аккаунтом на Facebook под каждой страницей введения личности после завершения теста на личность в 16personalities. Таким образом, мы ползаем все Facebook пользователя ид под каждой страницей знакомства с личностью, как основа правды для Facebook набор данных.

2. Коллекция контента, созданная пользователем. Затем мы начали загружать пользовательские данные из нескольких источников с помощью API Twitter REST и API Facebook GRAPH. Для защиты конфиденциальности пользователей информация пользователей была анонимизировала. В таблице представлена статистика нашего набора данных. Из таблицы видно, что INFP, INFJ, INTJ, ENTP, INTP являются Топ-5 распределенных личностных признаков как в Facebook, так и в Twitter. Из рисунка также видно, что существует долгосрочное распределение личностных признаков по источникам данных. Поэтому оценка должна проводиться по признакам отдельно, чтобы избежать проблемы оценки несбалансированного набора данных.

Мы представляем наши данные следующим образом.

1. Текстовые характеристики: Для каждого источника данных текстовые данные

каждого пользователя объединялись в документ, извлекались с помощью term frequency-inverse document frequency (TF-IDF) в матрицу document-term, а затем применялся скрытый семантический анализ (LSA) [6], как сообщалось в предыдущем исследовании для обеспечения эффективности аналогичного приложения для профилирования пользователей. Размерность разложенной функции установлена на 150, что показывает наилучшую производительность при поиске по сетке.

2. Особенности изображения: Мы автоматически привязываем каждую фотографию Instagram к 1000 ImageNet [1] концепции изображения, используя предварительно подготовленную модель ResNet-101 [8]. Затем мы суммировали прогнозируемые вероятности появления концепций для каждого пользователя и разделили полученный вектор на общее количество размещённых этим пользователем изображений. Всего мы извлекли 1000 возможностей изображения для каждого пользователя. Аналогично текстовой модальности, для разложения размерности на 200 к концепциям изображений был применён Принципиальный компонентный анализ (PCA).

### Результат

Чтобы ответить на вопросы нашего исследования, мы сравниваем производительность MMIPS, обученных с единым источником и комбинациями модальностей. Для оценки производительности системы MMIPS набор данных был равномерно разделен на обучающий набор и тестовый набор в соотношении 85:15. Для поддержания одинакового распределения личностных этикеток в реальном мире мы сохраняем одинаковое распределение образцов в тренировочном наборе и тестовом наборе. Для всесторонней оценки MMIPS мы выбрали базовые уровни следующим образом:

1. Обучающий классификатор с одной модальностью отдельно.
2. Early Fusion. Обучайте классификатор с ранним объединением, применяемым к текстовому элементу и элементу изображения, сначала объединяя их.
3. Early Fusion (PCA200). Сначала мы объединяем функцию текста и функцию изображения в новый обучающий набор и применяем к нему PCA, чтобы уменьшить размерность до 200.

В связи с несбалансированным распределением личностных этикеток и во избежание классификатора от предпочтения большинства классов, мы особенно используем макро F1 метрики, которые являются гармоническим средним между точностью и напоминанием, где среднее вычисляется на этикетке, а затем усредняется по всем этикеткам. Мы также использовали метрику коэффициента корреляции Мэттьюса (Mcor), так как она учитывает истинные и ложные положительные и отрицательные значения и обычно рассматривается как сбалансированная мера, которая может быть использована, даже если классы имеют очень разные размеры.

Во-первых, мы исследовали вклад различных способов профилирования личности и способности к интеграции. Интересное замечание исходит из комбинации источников данных, приведенных в таблице, где стекирование из одного источника вернуло прирост производительности по сравнению с нестекирующим подходом с одной модальностью при предсказании большинства личностных меток, за исключением метки SN в Twitter и Facebook. Более впечатляющий результат может быть получен от комбинации модальностей, где обучение с текстовыми и графическими данными, MMIPS превосходит классификатор, который обучался только на текстовых или графических данных и раннем подходе слияния, за исключением SN этикетки в Twitter. На основании результата можно сделать вывод, что визуальная модальность, без сомнения, важна для профилирования личности, а мультимодальный помогает улучшить производительность вывода личности. Это наблюдение можно объяснить,

предположительно, богатством визуальных данных для представления предпочтений пользователей, которые связаны с их личностными особенностями, и это было бы полезным дополнением к текстовым данным. В результате мы отвечаем на вопрос исследования "Помогают ли мультимодальные данные из аккаунта пользователя в социальных сетях улучшить работу вывода о личностных особенностях пользователя", подчеркивая решающую роль визуальных данных в задаче профилирования личности и рекомендуя их использование в исследованиях, в большей степени связанных с личностными особенностями.

Таблица

### Результат системы MMIPS

| Twitter Dataset              |           |           |           |           | Facebook Dataset        |           |           |           |           |
|------------------------------|-----------|-----------|-----------|-----------|-------------------------|-----------|-----------|-----------|-----------|
| Single Source (F1/Mcor)      |           |           |           |           | Single Source (F1/Mcor) |           |           |           |           |
| модель                       | EI        | SN        | TF        | JP        | модель                  | EI        | SN        | TF        | JP        |
| Text                         |           |           |           |           |                         |           |           |           |           |
| Xgboost                      | 0.80/0.62 | 0.48/0.08 | 0.59/0.23 | 0.61/0.22 | Xgboost                 | 0.63/0.26 | 0.59/0.22 | 0.55/0.16 | 0.58/0.17 |
| RF                           | 0.78/0.56 | 0.57/0.14 | 0.61/0.22 | 0.62/0.25 | RF                      | 0.61/0.21 | 0.62/0.26 | 0.59/0.17 | 0.58/0.16 |
| LGBM                         | 0.81/0.62 | 0.57/0.14 | 0.62/0.25 | 0.63/0.26 | LGBM                    | 0.63/0.25 | 0.62/0.24 | 0.57/0.13 | 0.58/0.16 |
| Image                        |           |           |           |           |                         |           |           |           |           |
| Xgboost                      | 0.46/0.01 | 0.47/0.03 | 0.51/0.01 | 0.54/0.09 | Xgboost                 | 0.59/0.19 | 0.47/0.05 | 0.49/0.07 | 0.56/0.14 |
| RF                           | 0.54/0.09 | 0.52/0.06 | 0.57/0.14 | 0.56/0.12 | RF                      | 0.59/0.18 | 0.54/0.08 | 0.57/0.15 | 0.57/0.14 |
| LGBM                         | 0.52/0.05 | 0.52/0.04 | 0.56/0.11 | 0.57/0.13 | LGBM                    | 0.59/0.18 | 0.55/0.11 | 0.57/0.14 | 0.56/0.12 |
| Early Fusion (PCA 200)       |           |           |           |           |                         |           |           |           |           |
| Xgboost                      | 0.79/0.61 | 0.48/0.07 | 0.59/0.23 | 0.60/0.20 | Xgboost                 | 0.57/0.14 | 0.47/0.04 | 0.50/0.07 | 0.55/0.10 |
| RF                           | 0.78/0.56 | 0.57/0.14 | 0.63/0.26 | 0.62/0.24 | RF                      | 0.60/0.21 | 0.54/0.08 | 0.58/0.16 | 0.56/0.13 |
| LGBM                         | 0.81/0.62 | 0.56/0.11 | 0.62/0.25 | 0.62/0.24 | LGBM                    | 0.59/0.18 | 0.52/0.04 | 0.56/0.12 | 0.56/0.13 |
| MMIPS -Single Source         |           |           |           |           |                         |           |           |           |           |
|                              | EI        | SN        | TF        | JP        |                         | EI        | SN        | TF        | JP        |
| Text                         | 0.81/0.62 | 0.5/0.12  | 0.62/0.26 | 0.63/0.28 | Text                    | 0.62/0.23 | 0.62/0.3  | 0.59/0.21 | 0.58/0.16 |
| Image                        | 0.53/0.11 | 0.47/0.05 | 0.57/0.16 | 0.59/0.17 | Image                   | 0.59/0.17 | 0.50/0.10 | 0.56/0.17 | 0.56/0.13 |
| MMIPS - Dual Source Stacking |           |           |           |           |                         |           |           |           |           |
|                              | EI        | SN        | TF        | JP        |                         | EI        | SN        | TF        | JP        |
| T+I                          | 0.82/0.61 | 0.50/0.12 | 0.63/0.26 | 0.64/0.28 | T+I                     | 0.64/0.28 | 0.63/0.31 | 0.60/0.23 | 0.59/0.17 |

Теперь давайте объясним результаты оценки из одного источника (таблица). Важно отметить, что в случае обучения из одного источника данных MMIPS, обученный из текстовой модели, работает лучше всего. Количественное превосходство может прояснить гегемонию текстовых данных над другими модальностями. Наконец, обученный визуальной модальности, MMIPS работает хуже, чем текстовая модальность. Одной из потенциальных причин этого является высокий уровень шума в загружаемых пользователем изображениях. Для сравнения расхождения с предыдущими исследованиями можно объяснить общностью концепций изображений ImageNet, которые могут быть полезны для общей задачи вывода демографического атрибута, но менее успешны для более узкой темы задачи профилирования личности.

### Вывод

В этой работе мы представили одно из первых исследований по сравнению производительности модели машинного обучения при профилировании черт личности между различными источниками (Facebook и Twitter) с предложенной нами моделью под названием MMIPS. Экспериментальные результаты по трем источникам показывают, что MMIPS способен повысить производительность при прогнозировании личностных характеристик пользователей путем агрегирования мультимодальных данных, генерируемых пользователями из их аккаунта в социальных сетях. Мы выпустили мультимодальный набор данных для профилирования личности, чтобы помочь исследовательскому сообществу в дальнейших исследованиях в похожих областях, таких как профилирование пользователей и система рекомендатель с улучшенными личностными характеристиками.

### Литература

1. Deng J. [и др.] ImageNet: A Large-Scale Hierarchical Image Database [Конференция] // CVPR09. 2009.
2. Farseev A. [и др.] Cross-domain recommendation via clustering on multi-layer graphs [Конференция] // Proceedings of the 40th International ACM SIGIR Conference on Research and Development in Information Retrieval. 2017. стр. 195–204.
3. Gjurković Matej и Šnajder Jan Reddit: A Gold Mine for Personality Prediction [Конференция] // Proceedings of the Second Workshop on Computational Modeling of People's Opinions, Personality, and Emotions in Social Media. - New Orleans, Louisiana, USA : Association for Computational Linguistics, 2018. стр. 87–97.
4. He K. [и др.] Deep Residual Learning for Image Recognition [Конференция] // 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR). 2016. стр. 770–778.
5. Kosinski Michal [и др.] Facebook as a Research Tool for the Social Sciences [Журнал] // The American psychologist. 9 2015 г. Т. 70. стр. 543–556.
6. Kumar K.N.P. и Gavrilova M.L. Personality Traits Classification on Twitter [Конференция] // 2019 16th IEEE International Conference on Advanced Video and Signal Based Surveillance (AVSS). 2019. стр. 1–8.
7. Mairesse F. [и др.] Using Linguistic Cues for the Automatic Recognition of Personality in Conversation and Text [Журнал] // J. Artif. Int. Res. El Segundo, CA, USA: AI Access Foundation, 11 2007 г. Т. 30. стр. 457–500. ISSN: 1076-9757.
8. Pennebaker J., Francis Martha и Booth Roger Linguistic inquiry and word count (LIWC) [Журнал]. 1. 1999 г..
9. Pennington Jeffrey, Socher Richard и Manning Christopher D. GloVe: Global Vectors for Word Representation [Конференция] // Empirical Methods in Natural Language Processing (EMNLP). 2014- стр. 1532–1543.
10. Tadesse Michael M. [и др.] Personality Predictions Based on User Behavior on the Facebook Social Media Platform [Журнал] // IEEE Access. 2018 г. Т. 6. стр. 61959–61969. ISSN: 21693536.
11. TF-IDF // Encyclopedia of Machine Learning / ред. Sammut Claude и Webb Geoffrey I. Boston : Springer US, 2010. ISBN: 978-0-387-30164-8.

---

**СОДЕРЖАНИЕ**

|   |          |
|---|----------|
| <b>НАПРАВЛЕНИЕ «ИНФОРМАЦИОННЫЕ ТЕХНОЛОГИИ И ПРОГРАММИРОВАНИЕ; ИНФОКОММУНИКАЦИОННЫЕ ТЕХНОЛОГИИ; ДИЗАЙН И УРБАНИСТИКА».....</b>   | <b>4</b> |
| <b>Антонов А.С., Иванов С.В.</b><br>ИНТЕГРАЦИЯ НЕЙРОННОЙ СЕТИ<br>ДЛЯ СЕМАНТИЧЕСКОЙ СЕГМЕНТАЦИИ<br>ОБЪЕКТОВ ПО ОБЛАКУ ТОЧЕК В КОМПЛЕКСНУЮ<br>СИСТЕМУ АНАЛИЗА И ОБРАБОТКИ ДАННЫХ.....                   | 5        |
| <b>Антропов Д.В.</b><br>ИССЛЕДОВАНИЕ ТРАНСПОРТНОГО<br>ОБЕСПЕЧЕНИЯ И РЕШЕНИЕ ПРОБЛЕМ<br>НА ОСНОВЕ ИНТЕЛЛЕКТУАЛЬНЫХ РЕШЕНИЙ.....  | 11       |
| <b>Безбородов А.К., Деева И.Ю.</b><br>АЛГОРИТМ ПОИСКА МЕСТОРОЖДЕНИЙ-АНАЛОГОВ<br>НА ОСНОВЕ МЕТОДОВ МАШИННОГО ОБУЧЕНИЯ.....   | 16       |
| <b>Бейлин М.Т.</b><br>ОБЗОРНЫЙ АНАЛИЗ МЕТОДОВ ОБРАБОТКИ<br>БОЛЬШИХ ДАННЫХ В УПРАВЛЯЮЩИХ СИСТЕМАХ<br>ПРОГНОЗИРОВАНИЯ ПОВЕДЕНИЯ РЫНКА.....  | 22       |
| <b>Бокарева А.А.</b><br>СБОР И АНАЛИЗ ДАННЫХ ПО СЕРДЕЧНО-СОСУДИСТЫМ<br>ЗАБОЛЕВАНИЯМ НА ФОНЕ COVID-19 ДЛЯ ВЫЯВЛЕНИЯ<br>ЗАКОНОМЕРНОСТЕЙ И ВЕРОЯТНЫХ ПРОГНОЗОВ НА ОСНОВЕ<br>СУЩЕСТВУЮЩЕЙ СТАТИСТИКИ..... | 26       |
| <b>Брагина Т.В., Землянова Е.И., Лаушкина А.А., Рослякова С.В.</b><br>ИССЛЕДОВАНИЕ ВЛИЯНИЯ РАЗЛИЧНЫХ<br>СВЕТОВЫХ РЕЖИМОВ НА СНИЖЕНИЕ<br>СТРЕССОВОГО СОСТОЯНИЯ ПОЛЬЗОВАТЕЛЯ.....                       | 31       |
| <b>Бубнова А.В.</b><br>МЕТОДЫ ПОИСКА НЕОДНОРОДНОСТЕЙ В СМЕШАННЫХ<br>ДАННЫХ НА ОСНОВЕ БАЙЕСОВСКИХ СЕТЕЙ.....   | 35       |
| <b>Вобликов С.Д., Андрусенко А.Ю.</b><br>ИЗУЧЕНИЕ РАЗЛИЧНЫХ РАСПИСАНИЙ ИЗМЕНЕНИЯ<br>СКОРОСТИ ОБУЧЕНИЯ НЕЙРОННОЙ СЕТИ, ИСПОЛЬЗУЕМОЙ<br>ДЛЯ АВТОМАТИЧЕСКОГО РАСПОЗНАВАНИЯ РЕЧИ.....                     | 40       |

|  |    |
|--|----|
| <b>Гец О.В.</b><br>ПРОЕКТИРОВАНИЕ ИТ-АРХИТЕКТУРЫ<br>НА НЕФТЕДОБЫВАЮЩЕМ ПРЕДПРИЯТИИ.....  | 46 |
| <b>Глухов В.Н.</b><br>ОБЗОР ТЕКУЩЕГО СОСТОЯНИЯ<br>МУЛЬТИВАРИАТИВНЫХ МЕТОДОВ ОТБОРА<br>ПРИЗНАКОВ НА ОСНОВЕ ТЕОРИИ ИНФОРМАЦИИ.....   | 50 |
| <b>Боченина К.О., Гороховатский Л.Ю., Суриков А.Г.</b><br>ТЕОРЕТИЧЕСКИЕ ПОДХОДЫ К ПРОБЛЕМЕ ИССЛЕДОВАНИЯ<br>ЭМОЦИЙ В КИБЕРПРОСТРАНСТВЕ И ИХ ПЕРЕСЕЧЕНИЕ<br>С КЛЮЧЕВЫМИ КОНЦЕПЦИЯМИ ЛИЧНОСТИ.....          | 54 |
| <b>Гусев А.Е., Волкова М.В., Астапов С.С., Кабаров В.И.</b><br>ИССЛЕДОВАНИЕ МЕТОДОВ ГОЛОСОВОЙ<br>БИОМЕТРИИ НА КОРОТКИХ ДЛИТЕЛЬНОСТЯХ.....  | 58 |
| <b>Двойникова А.А., Мамонтов Д.Ю., Карпов А.А.</b><br>АВТОМАТИЧЕСКОЕ ОПРЕДЕЛЕНИЕ<br>МОЦИОНАЛЬНОГО СОСТОЯНИЯ УЧАСТНИКОВ<br>ПРЕДМЕТНЫХ РАЗГОВОРОВ ПО ТРАНСКРИПЦИЯМ РЕЧИ.....                               | 63 |
| <b>Дудина Д.С.</b><br>АНАЛИЗ СПОСОБОВ И СРЕДСТВ ИЗВЛЕЧЕНИЯ<br>И ИНТЕРПРЕТАЦИИ ПАРАМЕТРОВ КЛАВИАТУРНОГО<br>ПОЧЕРКА ПОЛЬЗОВАТЕЛЯ.....  | 69 |
| <b>Евсеева Е.С.</b><br>ОЦЕНКА ПРИМЕНИМОСТИ МОДЕЛИ<br>COUNTNET ДЛЯ ПОДСЧЕТА ЧИСЛА ДИКТОРОВ.....   | 74 |
| <b>Емельянова В.И., Якимук И.В.</b><br>МЕТОД ОЦЕНКИ ПОТЕНЦИАЛА<br>НАБЕРЕЖНЫХ КАК ДОСУГОВЫХ ПРОСТРАНСТВ.....  | 78 |
| <b>Залуская В.С., Астапов С., Кабаров В.И.</b><br>ОСОБЕННОСТИ ЛЕКСИКИ В КОНТЕКСТЕ РЕШЕНИЯ<br>ЗАДАЧИ ОПРЕДЕЛЕНИЯ ТОЧЕК СМЕНЫ ДИКТОРА.....   | 82 |
| <b>Землянова Е.И., Брагина Т.В., Лаушкина А.А., Рослякова С.В.</b><br>ИССЛЕДОВАНИЕ ВЛИЯНИЯ РАЗЛИЧНЫХ СВЕТОВЫХ<br>РЕЖИМОВ НА РАЗРЕШЕНИЕ КОНФЛИКТНОЙ СИТУАЦИИ<br>В МНОГОПОЛЬЗОВАТЕЛЬСКОМ ПРОСТРАНСТВЕ..... | 87 |
| <b>Исмагилова А.Ф.</b><br>АНАЛИЗ СУЩЕСТВУЮЩИХ СРЕДСТВ<br>ОЦЕНИВАНИЯ УДОВЛЕТВОРЕННОСТИ<br>СТУДЕНТОВ ЗАНЯТИЕМ В ДИСТАНЦИОННОМ<br>ФОРМАТЕ ПОСРЕДСТВОМ ВИДЕОКОНФЕРЕНЦСВЯЗИ.....                              | 91 |

|   |     |
|---|-----|
| <b>Капранова В.О., Астапов С.С., Кабаров В.И.</b><br>СРАВНЕНИЕ МЕТОДОВ ВЫДЕЛЕНИЯ РЕПЛИК ЦЕЛЕВОГО<br>ДИКТОРА В ЗАПИСИ РЕЧИ НЕСКОЛЬКИХ ДИКТОРОВ.....  | 97  |
| <b>Кирилова С.А.</b><br>ПРИМЕНЕНИЕ НЕЙРОННЫХ СЕТЕЙ В ПОСТРОЕНИИ<br>ИНДИВИДУАЛЬНОЙ ТРАЕКТОРИИ ОБУЧЕНИЯ.....  | 102 |
| <b>Кшенин А.Д.</b><br>ИДЕНТИФИКАЦИЯ И ИССЛЕДОВАНИЕ<br>ИМИТАЦИОННЫХ МОДЕЛЕЙ БИЗНЕС-ПРОЦЕССОВ<br>В КРУПНОЙ КОМПАНИИ ПО ДАННЫМ КОРПОРАТИВНЫХ<br>ИНФОРМАЦИОННЫХ СИСТЕМ.....   | 105 |
| <b>Ле Ань Ту</b><br>СРАВНИТЕЛЬНЫЙ МЕТОД ОБНАРУЖЕНИЯ<br>ДВИЖУЩИХСЯ НАЗЕМНЫХ ЦЕЛЕЙ.....   | 109 |
| <b>Липатова Д.С.</b><br>ВНЕДРЕНИЕ ЭЛЕКТРОННЫХ<br>ГОСУДАРСТВЕННЫХ СОЦИАЛЬНЫХ УСЛУГ<br>В ПРОАКТИВНОМ ФОРМАТЕ В САНКТ-ПЕТЕРБУРГЕ.....  | 114 |
| <b>Логунов А.А.</b><br>СРАВНЕНИЕ МЕТОДОВ ПОТОКОВОЙ ДИАРИЗАЦИИ.....  | 118 |
| <b>Мамаев Н.К.</b><br>ОЦЕНКА НЕУВЕРЕННОСТИ НЕЙРОННЫХ<br>МОДЕЛЕЙ ДЛЯ ТЕКСТОВОЙ КЛАССИФИКАЦИИ.....  | 122 |
| <b>Маркова Ю.С.</b><br>АНАЛИЗ И ИССЛЕДОВАНИЕ БЫСТРОГО<br>ВЫПОЛНЕНИЯ СЦЕНАРИЕВ ТЕСТИРОВАНИЯ<br>ИНФОКОММУНИКАЦИОННОЙ СИСТЕМЫ.....   | 126 |
| <b>Маслюхин С.М.</b><br>ОПТИМИЗАЦИЯ ГЛУБОКИХ НЕЙРОННЫХ<br>СЕТЕЙ НА БАЗЕ ТРАНСФОРМЕРОВ.....  | 130 |
| <b>Машина Е.А.</b><br>УЧЕТ ОСОБЕННОСТЕЙ ПРЕОБРАЗОВАНИЯ<br>УЗКОСПЕЦИАЛИЗИРОВАННЫХ ПРОФЕССИОНАЛЬНЫХ<br>ТЕКСТОВ К ПОНЯТИЙНОМУ РЯДУ, ПОНЯТНОМУ<br>НЕКВАЛИФИЦИРОВАННЫМ ПОЛЬЗОВАТЕЛЯМ,<br>ПРИ ВЫБОРЕ ТЕХНОЛОГИЙ МАШИННОГО ПЕРЕВОДА,<br>ОСУЩЕСТВЛЯЕМОГО ВНУТРИ ОДНОГО ЯЗЫКА..... | 134 |
| <b>Мишина М.Э., Гурина Е.П.</b><br>АНАЛИЗ ИНСТРУМЕНТОВ ПОИСКА<br>ОПТИМАЛЬНОГО МАРШРУТА<br>В ТЕЛЕКОММУНИКАЦИОННОЙ СЕТИ В СРЕДЕ ARCGIS.....   | 140 |

|  |     |
|--|-----|
| <b>Назаренко У.К., Козырева А.И.</b><br>ИНТЕГРАЦИЯ ПЛАТФОРМ СОЦИАЛЬНЫХ МЕДИА<br>И СПЕЦИАЛИЗИРОВАННЫХ ВЕБ-РЕСУРСОВ<br>ДЛЯ ЭФФЕКТИВНОГО ИСПОЛЬЗОВАНИЯ<br>ВЫСОКОТЕХНОЛОГИЧНОЙ МЕДИЦИНСКОЙ ИНФОРМАЦИИ..... | 145 |
| <b>Беген П.Н., Низомутдинов Б.А.</b><br>РАЗРАБОТКА МЕТОДА ИЗВЛЕЧЕНИЯ АДРЕСА ИЗ ТЕКСТА<br>СООБЩЕНИЙ О ПРОИСШЕСТВИЯХ В СОЦИАЛЬНЫХ<br>СЕТЯХ ДЛЯ ФОРМИРОВАНИЯ ГЕОПРИВЯЗАННЫХ ДАННЫХ.....                   | 152 |
| <b>Низомутдинов Б.А., Углова А.Б.</b><br>МЕТОД МОНИТОРИНГА ПОТРЕБНОСТЕЙ<br>ЖИТЕЛЕЙ В РАЗВИТИИ ГОРОДСКОЙ ИНФРАСТРУКТУРЫ.....  | 156 |
| <b>Николаева К.А.</b><br>МОДЕЛИРОВАНИЕ ИНФОРМАЦИОННЫХ<br>ПРОЦЕССОВ В СЛОЖНЫХ РАСПРЕДЕЛЕННЫХ<br>ТЕЛЕМЕДИЦИНСКИХ СИСТЕМАХ.....   | 160 |
| <b>Огирь К.С.</b><br>ПЕРСПЕКТИВЫ ИСПОЛЬЗОВАНИЯ МАШИННОГО<br>ОБУЧЕНИЯ ПРИ АНАЛИЗЕ ПРОГРАММНОГО КОДА.....  | 163 |
| <b>Панфилов Г.О.</b><br>СРАВНИТЕЛЬНЫЙ АНАЛИЗ ПРОГРАММНОГО<br>ОБЕСПЕЧЕНИЯ ДЛЯ АВТОМАТИЗИРОВАННОГО<br>СБОРА ИНФОРМАЦИИ ИЗ СОЦИАЛЬНЫХ СЕТЕЙ.....  | 168 |
| <b>Рюмина Е.В., Дресвянский Д.В.</b><br>ВЛИЯНИЕ МЕТОДОВ СНИЖЕНИЯ<br>СКОРОСТИ НА ЭФФЕКТИВНОСТЬ СИСТЕМ<br>РАСПОЗНАВАНИЯ ЭМОЦИЙ ПО МИМИКЕ ЛИЦА.....   | 174 |
| <b>Скрипачев И.С.</b><br>РАЗРАБОТКА МЕТОДА ПЛАНИРОВАНИЯ<br>ОПТИМАЛЬНОГО РАЗМЕЩЕНИЯ<br>ОБЪЕКТОВ НАРУЖНОГО ОСВЕЩЕНИЯ.....  | 181 |
| <b>Соловьева В.В.</b><br>АНАЛИЗ ФАКТОРОВ, ВЛИЯЮЩИХ НА СТОИМОСТЬ<br>ЭКОНОМИИ ВРЕМЕНИ В ПУТИ ЖИТЕЛЕЙ ГОРОДА.....   | 187 |
| <b>Сома Гедеш Мануэл</b><br>АНАЛИЗ СУЩЕСТВУЮЩИХ СРЕДСТВ<br>БИМОДАЛЬНОГО ОПРЕДЕЛЕНИЯ ЭМОЦИЙ<br>В ХОДЕ ЗАНЯТИЙ В ДИСТАНЦИОННОМ ФОРМАТЕ.....  | 192 |
| <b>Ставинова Е.А., Чунаев П.В.</b><br>КЛАССИФИКАЦИЯ КЛИЕНТОВ БАНКА СОГЛАСНО<br>ИНДИВИДУАЛЬНОЙ ПРЕДСКАЗУЕМОСТИ ИХ ПОВЕДЕНИЯ.....  | 198 |

|   |     |
|---|-----|
| <b>Тимофеева Е.П.</b><br>ОПРЕДЕЛЕНИЕ ПЕРЕКРЫТИЙ<br>В РЕЧИ С ПОМОЩЬЮ МОДЕЛИ COUNTNET.....  | 205 |
| <b>Тихонова Т.А., Хватов А.А.</b><br>ЧИСЛЕННОЕ РЕШЕНИЕ ДИФФЕРЕНЦИАЛЬНЫХ<br>УРАВНЕНИЙ В ЧАСТНЫХ ПРОИЗВОДНЫХ<br>С ПОМОЩЬЮ МЕТОДОВ ОПТИМИЗАЦИИ.....  | 209 |
| <b>Хрульков А.А., Терентьев Р.А.</b><br>КОМПЬЮТЕРНОЕ ЗРЕНИЕ ДЛЯ АНАЛИЗА<br>И ПОДСЧЕТА ГОРОДСКОГО ТРАФИКА.....   | 213 |
| <b>Чурякова Т.А., Якимук И.В.</b><br>МЕТОД ОЦЕНКИ ПОТЕНЦИАЛА ЗЕЛЕННЫХ<br>ЗОН ДЛЯ СОЗДАНИЯ ДИСПЕРСНЫХ ПАРКОВ.....  | 216 |
| <b>Щепина Е.К., Егорова Е.В., Федотов П.А.</b><br>АНАЛИЗ ИНТЕРЕСОВ И КУЛЬТУРНЫХ ПРЕДПОЧТЕНИЙ<br>ПОЛЬЗОВАТЕЛЕЙ СОЦИАЛЬНЫХ СИСТЕМ<br>НА ПРИМЕРЕ ИХ МУЗЫКАЛЬНЫХ ПОДПИСОК<br>И ИХ ВЗАИМОСВЯЗЬ С ПСИХОТИПОМ..... | 223 |
| <b>Эльдиб П.Ю.</b><br>ФОРМИРОВАНИЕ ФУНКЦИОНАЛЬНЫХ<br>ТРЕБОВАНИЙ К ГЕОИНФОРМАЦИОННЫМ<br>СИСТЕМАМ ОБЩЕСТВЕННОГО УЧАСТИЯ.....  | 230 |
| <b>Ян Ци</b><br>TOWARDS MULTI-VIEW MULTI-SOURCE LEARNING FROM SOCIAL<br>MULTIMEDIA FOR HUMAN PERSONALITY PROFILING.....   | 235 |

**АЛЬМАНАХ  
НАУЧНЫХ РАБОТ  
МОЛОДЫХ УЧЁНЫХ  
Университета ИТМО**

Том 3

Редакционно-издательский отдел Университета ИТМО

Зав. РИО

Дизайн обложки

Вёрстка

Подписано к печати 26.10.2021

Заказ № 4526 от 26.10.2021

Тираж 100 экз.

Н.Ф. Гусарова

Н.А. Потехина

Я.Я. Платунова

Печатается в авторской редакции

---

Отпечатано: Учреждение «Университетские коммуникации»  
199034, Санкт-Петербург, В.О., Биржевая линия, 16

ISBN 978-5-7577-0649-8



9 785757 706498

Редакционно-издательский отдел  
Университета ИТМО

197101, Санкт-Петербург, Кронверкский пр., 49